

Open Research Online

The Open University's repository of research publications and other research outputs

Parental discipline style : an investigation into the validity of parents' reports

Thesis

How to cite:

Bellhouse, Margaret Ruth (2001). Parental discipline style : an investigation into the validity of parents' reports. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2001 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000d468>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

**Parental Discipline Style:
an Investigation into the Validity of Parents' Reports**

by

Margaret Ruth Bellhouse, BA, M.Sc

July, 2001

SUBMITTED IN PART FULFILMENT OF THE
OU/BPS DOCTORATE IN CLINICAL PSYCHOLOGY,
ISIS EDUCATION CENTRE, WARNEFORD HOSPITAL, OXFORD

AWARD DATE: 11 September 2001

ABSTRACT

Parenting is thought to play an important role in the development and maintenance of children's behaviour difficulties. Research involving the assessment of parent-child interactions has contributed greatly to the current understanding in this area. The present study sought to assess the concurrent validity of a popular self-report measure of parental discipline style, by examining the relationship between parents' reports of these aspects of their parenting behaviour and objective observations of the same behaviour. This represents one of only a few investigations of this kind in relation to parenting behaviour.

Parents of children who experienced significant behavioural difficulties completed questionnaires about their parenting style, depression, sense of competence in their parenting and their children's behaviour problems. These parents were also observed in interactions with their children at home. An observational coding system was developed for the study. Analysis of the observational system indicated that it was both reliable and valid. However, there was no relationship between observed and parent-reported discipline styles. Comparisons between parent-reported discipline style and other questionnaire factors led to mixed results. Taken together, these results suggested that the self-report measure of parenting style did not have good concurrent validity, and this finding contrasted with those reported by the questionnaire's developers.

These results were interpreted with reference to the design of the questionnaire, respondent-related factors, deficits in parental monitoring skills, and biases in response to the questionnaire. These results have implications for therapeutic intervention and for future research involving parental self-report measures of discipline, and these are discussed.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Frances Gardner, for all her helpful advice and invaluable proof reading. My thanks must also go to the research team in general (Frances, Jenny Burton and Sue Kirkpatrick) for making me an honorary member of the group for the last year. They may not realise it now, but one day they'll miss my regular and multiple E-Mails - or maybe not! For my part, I'll miss putting the world to rights with Jenny... I wish the team good luck with the remainder of the project.

I would also like to thank Myra Cooper and Paul Griffiths for their great advice and support throughout, and Helen Aspland for relieving the panic more than once!

I have found all members of my family to be very supportive as always, in terms of practical advice, emotional support, and the appearance of being interested in what I was doing however many times I told them about it! Too many to mention everyone by name – you know who you are! - though I must pay tribute to my dear grandfather, Francis, and my inspirational grandmother, Helen: both gone but not forgotten by any means.

Finally, I would like to thank Phil, but adequate words escape me. Suffice to say that his support has also been fantastic in all respects. I'm sorry for all the weekends sacrificed for this work, but I'm looking forward to making up for this in the future.

CONTENTS

Abstract		Page 1
Introduction		Page 1
	Background	Page 1
	Assessment of Parenting	Page 7
	Multimethod Parenting Research	Page 23
	Conclusions	Page 28
	The Present Study	Page 28
Method		Page 30
	Participants	Page 30
	Design	Page 30
	Measures	Page 31
	Procedure	Page 40
	Ethical Issues	Page 41
Results		Page 43
	Description of Sample	Page 43
	Behaviour Coding System Reliability Analysis	Page 45
	Analysis of the Concurrent Validity of the Parenting Scale	Page 51
Discussion		Page 59
	The Concurrent Validity of the Parenting Scale	Page 61
	The Validity of the Observational Coding System	Page 64
	Interpretation of the Results	Page 68
	Clinical and Research Implications	Page 72
	Limitations	Page 74
	Conclusions	Page 77
References		
Appendices		

INTRODUCTION

BACKGROUND

Behavioural Difficulties in Children

Behavioural difficulties described as externalising, such as defiance, aggression, and tantrums are common in young children, and this finding is not restricted to Western culture. Researchers across the world have estimated that between ten and fifteen percent of pre-school children have mild to moderate externalising behavioural difficulties, as defined by their scores on behavioural checklist measures (Campbell, 1995). Though such behaviour may be generally annoying or difficult for the family as they develop, many children are acting in ways that are related to their developmental stage. For these children, such behaviours are best viewed as age-appropriate and likely to cease as they grow older. For example, children aged between two and four years often exhibit defiance and discipline problems result, as they try to gain autonomy yet look for certain limits from caregivers (Campbell, 1995).

However, half of all children with externalising behavioural difficulties go on to develop more severe behavioural and emotional problems throughout life, including conduct disorder (Gardner, Sonuga-Burke & Sayal, 1999; Campbell, 1995; Webster-Stratton & Herbert, 1994). The diagnostic criteria for conduct disorder are

-
- A. A repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated, as manifested by the presence of three (or more) of the following criteria in the last 12 months, with at least one criterion present in the past 6 months:

Aggression to people and animals

- (1) often bullies, threatens, or intimidates others
- (2) often initiates physical fights
- (3) has used a weapon that can cause serious physical harm to others (e.g., a bat, brick, broken bottle, knife, gun)
- (4) has been physically cruel to people
- (5) has been physically cruel to animals
- (6) has stolen while confronting a victim (e.g., mugging, purse snatching, extortion, armed robbery)
- (7) has forced someone into sexual activity

Destruction of property

- (8) has deliberately engaged in fire setting with the intention of causing serious damage
- (9) has deliberately destroyed others' property (other than by fire setting)

Deceitfulness or theft

- (10) has broken into someone else's house, building, or car
- (11) often lies to obtain goods or favours or to avoid obligations (i.e., 'cons' others)
- (12) has stolen items of nontrivial value without confronting a victim (e.g., shoplifting, but without breaking and entering; forgery)

Serious violations of rules

- (13) often stays out at night despite parental prohibitions, beginning before age 13 years
- (14) has run away from home overnight at least twice while living in parental or parental surrogate home (or once without returning for lengthy period)
- (15) often truant from school, beginning before age 13 years

- B. The disturbance in behaviour causes clinically significant impairment in social, academic, or occupational functioning.

- C. If the individual is age 18 or over, criteria are not met for Antisocial Personality Disorder.
-

Table 1: DSM-IV Diagnostic Criteria for Conduct Disorder

outlined in Table 1 (DSM-IV, 1994).

Severe behavioural problems, which have been estimated to be present in approximately ten percent of the population, are usually the most common reason for

the referral of children to health services (Robins, 1991) and are more than twice as common as emotional disorders (Carr, 1999). Such behavioural problems are considered difficult to treat, and are costly to society (Gardner *et al.*, 1999).

Aetiological Concerns

Given the poor prognosis of externalising behaviours and the associated growing costs to society, the question of what factors are involved in the development and the maintenance of externalising behavioural problems is an important one, the answer to which would have implications for treatment. Many factors have been highlighted as playing important roles in the development of such difficulties, and include, among others, genetic predisposition, elevated testosterone levels, 'difficult' child temperament, insecure attachments, social skills and problem solving skills deficit, and disorganised family interactions (Carr, 1999). It is likely that a number of these factors act together to predispose a child to externalising behavioural difficulties. A growing body of research provides support for the observation that deficient parenting skills play a central role in both the development and maintenance of child behavioural difficulties. It seems highly possible that ineffective parenting acts as a mediator between other vulnerability factors and the externalising behavioural problems (Webster-Stratton & Herbert, 1994; Reid & Patterson, 1989).

In particular, it appears that inconsistent, harsh or excessively lax discipline practices on the part of parents are associated with delinquency and aggression in the child (O'Leary, 1995; Campbell, 1995). This finding has been greatly replicated by many researchers. Parental behaviours or parent-child relationship characteristics

found to be most influential in terms of the eventual development of child behavioural problems have been identified by researchers. These include: use of indirect commands; lack of enforcement; demonstration of affection during discipline episodes; lack of supervision; inconsistent limit setting and use of discipline; ineffective monitoring, and poor responsiveness. Excessive criticism, lack of warmth and positive interaction with the child, and regular use of physical punishment or shouting at the child have also been found to play a significant role in the development of child behaviour problems (Carr, 1999; Gardner, 1998, 1994, 1989, 1987; O'Leary, 1995; Webster-Stratton & Herbert, 1994; Arnold, O'Leary, Wolff & Acker, 1993; Vuchinich, Bank & Patterson, 1992; Patterson & Bank, 1986; Patterson, 1982).

The way in which these parental behaviours and parent-child relationship factors lead to the development of child externalising behaviours has also been debated. Patterson, an early pioneer in parenting research, proposed a highly influential social-learning and developmental 'coercion theory' of children's externalising behavioural difficulties. This theory states that the child's problematic behaviour is shaped and reinforced through thousands of conflict events within the family. The child learns through these conflicts over time that he or she can escape parental criticism through an escalation of the problematic behaviour, as the parent develops a tendency to withdraw from the conflict, and thus negatively reinforces the problematic behaviour on the child's part. The parent's behaviour is also negatively reinforced, through the temporary cessation of the conflict as the child wins. Alternatively, the parent may try to coax the child to behave in the way the parent wants to during times of conflict, in ways that ultimately benefit the child (e.g., use of

a bribe). This approach positively reinforces the child's behaviour, as well as that of the parent if the child complies with their requests. Over time and through repetition, both the child's and the parent's behaviours as described above become habitual, and the child's behaviour is generalised outside the home (Gardner, 1998; Webster-Stratton & Herbert, 1994; Reid & Patterson, 1989; Patterson, 1982). Patterson's theory has gained much empirical support (Serketich & Dumas, 1996).

Other not incompatible ways in which ineffective parenting appears to precipitate the development of children's externalising behavioural difficulties include poor infant-parent attachment influenced by lack of warmth and consistency in their relationship, as well as a lack of modelling of effective problem-solving and interpersonal skills (Carr, 1999).

Interventions

There may be many other factors influencing a parent's style of parenting (i.e., the particular way in which parents raise their children, with respect to practices such as consistency, use of boundaries and discipline methods, and displays of emotion). Such parent-related factors include: the parent's own experiences of being parented; parental mental health and emotional reactivity; the quality of the parents' relationship with one another; beliefs about parenting methods and abilities; and parental attribution style. Other influential factors include: economic status; child temperament and gender; external stress; and available support (Bradley & Corwyn, 1999; Smith & O'Leary, 1995; Webster-Stratton & Herbert, 1994; Simons, Whitbeck, Conger & Chyi-In, 1991).

Individual, family, group, and wider system interventions have been aimed at these and other factors influencing parental discipline style. Although such interventions may form a supplement to parenting skills approaches for improved results, many clinicians and researchers have instead focussed on parenting skill programmes alone. Research to date has led to the identification of parenting methods thought to be effective in promoting healthy development of a wide range of skills and positive behaviour in the child. Parenting skills programmes, based mainly on behavioural principles, aim to change the child's behaviour by teaching parents to use more effective parenting techniques, and thus to change the interpersonal antecedents and consequences that are eliciting and maintaining the child's negative behaviours (Webster-Stratton & Herbert, 1994). These programmes have consistently been found to be effective in the short- to medium- term, and more effective than eclectic programmes or waiting list control, but research is lacking with regards to longer-term outcome (Serketich & Dumas, 1996; Webster-Stratton & Herbert, 1994).

A particularly well-researched approach is based on social-learning and conditioning behavioural principles, although the approach is primarily a cognitive-behavioural one (Webster-Stratton & Herbert, 1994; Webster-Stratton, 1992, 1984, 1982, 1981). Parenting skills thought to be paramount in promoting positive behaviour in the child, and thus form the focus during this type of programme include: appropriate use of differential attention and other reinforcements; shaping; modelling; problem solving; parental emotion management; time-out; clear limit setting; and natural and logical consequences. These skills are discussed within a context of positive relationship emphasis, consistency and responsiveness. (Webster-Stratton & Herbert, 1994; Webster-Stratton, Hollinsworth & Kolpacoff, 1989).

ASSESSMENT OF PARENTING

Importance of Assessment of Parenting Skills

It has been argued above that parenting style appears to play an important part in the aetiology of children's externalising behavioural difficulties. Research involving the assessment of the parent-child interaction, especially using observational methods, has contributed greatly to the current understanding of how such behavioural difficulties develop and are maintained (Gardner, 1998). In addition, assessment of parenting skills is important in terms of the identification of parents whose discipline strategies are ineffective, with a view to appropriate intervention (Arnold *et al.*, 1993). Further, the aim of parenting skills programmes, as stated above, is to change children's behaviour via the development of parenting skills. Hence, it would seem that assessment of parenting skills both before and after a parenting skills programme should be an extremely important part of the evaluation of such a programme, and this does appear to be an increasingly common practice (Gardner, 2000). Therefore, the assessment of parenting skills may be considered a valuable part of both exploratory and evaluative research in this field. The vast majority of the research regarding parenting has concentrated on the interactions between parents and children up to the age of adolescence, and this trend is reflected in the studies reviewed below.

The two main methods of assessment of parenting skills used to date have been home and laboratory observation, and parental self-report, via printed measures or interview. Evidence has accrued to suggest that use of systematic observation

methods used in combination with self-report measures can be both valid and reliable in the assessment of parent-child interactions. However, correlations between observed and self-report data have typically been low to moderate, which suggests that a large amount of unique information is provided by each method (Gardner, 2000). These two approaches may be compatible. However, the choice of method depends greatly on the precise purpose of the assessment, and the match of assessment method to the aims of the investigation has implications for the validity of the findings. In the search for the most appropriate and efficient methods of assessment that yield sufficient levels of detail in relation to the research question, there are compelling arguments for their use individually.

It is to these two assessment methods that the remainder of the report will turn, with the ultimate aim of exploring the degree to which the information gained through these different methods is unique or shared. As self-report measures, interview techniques share some of the features and therefore the advantages of questionnaires. However, they also share some of the features of observational methods, particularly the role of interviewer interpretation and coding of response (e.g., Webster-Stratton & Spitzer, 1991).

In embarking on an investigation of alternative methods of assessment to observation that would not have the disadvantages of the latter method, it would seem appropriate to look to its opposite rather than methods with which it has key features in common. Questionnaires, as efficient and convenient self-report measures, may be considered opposite to observational methods for many reasons, and so for the remainder of this report, parenting research will be discussed with reference to

observation and self-report questionnaire methods only. The advantages and disadvantages of these two methods employed in isolation are discussed below, and summarised in Table 2. For illustrative purposes, examples of research using these methods separately will also be described.

Observational Assessment of Parenting Skills

Observation as a method of assessment can take many forms. It may be relatively informal participant observation, where the observer joins the group of individuals he or she is observing, without his or her role necessarily being known to the other group members. Observation may also be more formal, the most extreme form being complete observation where the observer does not participate at all in the interactions under observation, and the observer's precise role is often not communicated to those being observed. The formal method of observation is more appropriate in parenting skills assessment, where the observer may aim to minimise the effect of their presence on the interactions between parent and child. The nature of the observation for this purpose may be predetermined, as observers decide in advance that they are systematically assessing a range of skills identified previously but not communicated to the parent, and this avoids potential reporter bias on the parents' part. Alternatively, observation may be more casual when conducting exploratory research (Gardner, 2000; Clark-Carter, 1997). Formal observational

	Advantages	Disadvantages
Observation	<ul style="list-style-type: none"> • Versatile, flexible and tailored • Predetermined observation variables • Not susceptible to reporter bias • Data can be analysed quantitatively, using coding systems • Many coding systems found to be valid through longitudinal research • Can be used to collect detailed information in sequence concerning moment-to-moment interactions • Can be used to examine rates and proportions of behaviours • Thought more sensitive to treatment change than self-report measures • Can be conducted in natural and laboratory settings 	<ul style="list-style-type: none"> • Limited generalisation of findings to different settings and tasks • Data collection, coding and analysis time consuming, and often requires training • May be difficult to attain good reliability • Expensive process • Not appropriate for screening • Difficult to carry out in settings where the parent may be stressed or anxious • Small data sets • Low occurrence of some behaviours make them difficult to observe • Potential low stability of data • Somewhat burdensome to families • Participant anonymity not possible
Self-report	<ul style="list-style-type: none"> • Relatively cheap • Development process comparatively straightforward • Easy to complete and administer • Highly convenient • Appropriate for a range of purposes • Can easily use combination of measures • Training usually not required • High availability of range of measures • Potentially more appealing for parents • Allows a wide range of behaviours to be assessed • Greater data set • Limited opportunities for variability in researcher interpretation • Participant anonymity possible 	<ul style="list-style-type: none"> • Vulnerable to participant biases • Open to interpretation • Range of potentially confounding factors • Limited amount of response detail • Reliant on memory, understanding, insight and literacy of participant • Limited sensitivity to small changes • Low response rates likely if by post

Table 2: Advantages and Disadvantages of Observation and Self-Report Methods of Parenting Skills Assessment

methods are often accompanied by careful coding systems, such that the observation data can be analysed in quantitative form. Many complex coding systems exist for the observation of a wide range of parent and child behaviours, and longitudinal research involving these systems have found many of them to have both construct and predictive validity (Gardner, 2000).

Observational methods have the advantage that they may be used to gather fine detail in sequences regarding the matter under study, and this is especially true where the events are recorded by video equipment for future scrutiny. Hence, valuable information about multiple parent-child interactions, including behavioural triggers and reinforcement, collected in a way that reflects proportional rates of behaviours can be gained in this manner. Some argue that no other method can lead to the same level of detail in this respect (Gardner, 2000; 1998; 1997). Hence, observational methods are considered by some to be more sensitive to measuring any level of treatment change than self-report measures (Taylor, Schmidt, Pepler & Hodgkins, 1998). Observation can be conducted in natural settings such as the home, where much of the difficult behaviour of children is reported to occur, as well as in the laboratory. However, parent-child interactions observed in such artificial settings, whilst useful in terms of clinical work, are not necessarily representative of those occurring at home (Gardner, 1997).

It is extremely important that the behaviours observed are those that would occur naturally, as the extent to which this is true is closely linked to the external validity of the observation exercise. In an attempt to establish whether observed behaviours are representative of those that would occur naturally, researchers have

focused on participants' reactivity in relation to being observed. Studies involving the manipulation of observer obtrusiveness, and those including an assessment of presumed habituation effects over time, have been reassuring, as it seems that reactions to being observed do not significantly alter the interactions under study beyond the first few minutes. Hence, participant reactivity to being observed does not appear to compromise the construct validity of the method, although researchers should not become complacent in assuming that this is true in all cases, and it may be useful for the participants to meet informally with the observer before the formal observation sessions take place (Gardner, 2000;1997). Little is known about the impact of video recording in relation to the interactions under observation. Although more research regarding this issue is required, many assume that since a video recorder may seem less obtrusive than other methods, reactivity to the use of such equipment should not significantly affect the validity of the observation (Gardner, 2000; 1997).

It is quite common for researchers employing observation as a method of assessment of parenting skills to set a series of structured tasks for the caregivers and children to undertake, such as specific parent-child games, or the parent asking the child to clear up after a game. The use of such tasks may lead to more of the behaviours of interest occurring in the time of the observation. A further advantage to this approach is that the reliability of the observations tends to increase as the range of possible situational influences of the behaviour of interest decrease (Gardner, 2000). A few researchers have sought to establish whether the completion of such tasks would involve interactions comparable to those observed when no such tasks were imposed on the participants. The results generally suggest that researchers should

exercise caution when assuming that behaviours elicited by structured tasks are representative of those that occur in more natural circumstances, and are predictive of child behaviour in general. However, use of structured tasks as well as more naturalistic observation may still prove useful, especially when evaluating intervention outcome (Gardner, 2000).

It is clear that observation as a method of assessment of parenting skills has many advantages, and observational work has dominated this area of research for several decades to date, allowing investigators to study parent-child interactions in minute detail (Gardner, 1998; Robinson & Eyberg, 1981). However, observation as a research method is also accompanied by difficulties, such that it may not always be the automatic method of choice.

Although observation may allow a researcher to collect very detailed data systematically, this approach is very time-consuming. Following the planning stage which, like that in any study, can take a considerable amount of time, the data collection itself often takes an hour or more per participant for every visit: this is usually necessary in order to collect a sufficient amount of data (Gardner, 1997). Coding systems, used to transform the data to quantitative form, can be labour-intensive in their development as attaining reliability across coders can be a lengthy a difficult process. Once developed, such coding systems are often complex and require substantial training for its users. Hence, the use of observation as an assessment method is usually expensive, in terms of time, labour and financial cost (Dreyer, Mendelsohn & Tamis-LeMonda, 1996; Arnold *et al.*, 1993). In addition, observation methods do not lend themselves well to screening for potential clients

who are at risk but have not yet sought help, as it would not be efficient to use such time-consuming and expensive approaches in this way.

Although observation often takes place over at least an hour, it is unlikely that all behaviours of interest will occur during such a discrete time-period, especially more extreme behaviours that usually occur relatively infrequently (e.g., severe physical discipline). This, in combination with the time, effort and expense of the operation, inevitably leads to small data sets that potentially are of low stability (Gardner, 2000). When inter-observer reliability is low for reasons such as low stability of the data set, then the validity of the coding system is also limited.

In comparison to self-report measures, observation may appear burdensome to participants, and potential participants may not consent to such research for this reason. This may be especially so if the research involves investigators visiting the home, and it is difficult to carry out observation in settings where the parent may be stressed or anxious, with results likely to be limited in terms of generalisation to other settings (Gardner, 2000).

One final limitation relating to observational assessment is the impossibility of maintaining anonymity of those being assessed: at the very least, the observer will know the identity of the family members being observed, even if other members of the research team are blind to this fact. Some potential research participants may be lost for this reason.

Examples of Observation-Based Parenting Research

As mentioned previously, much of what is now known about the relationship between parenting practices and children's behaviour has been gained through observational research. For example, Patterson (1982) conducted detailed observations of parents and children in the home. This work led to the finding that parents of children with behavioural difficulties tended to be more harsh, punitive, erratic and inconsistent in their responses to the child compared to those whose children did not exhibit behavioural problems (*cf.* Gardner, 1998). These and other findings led Patterson to propose his persuasive coercion theory, outlined above.

Similarly, Wahler and Sansbury (1990) observed and videotaped 33 mothers together with their children, who had conduct disorders. The researchers also invited the mothers to comment on their children's behaviour when viewing themselves on video. These investigators found that the mothers tended to define their children's aversive behaviour using non-specific yet personalised terms, were highly inconsistent and conservative in their coding of this behaviour, and responded indiscriminately to their children's prosocial and aversive behaviour. The researchers interpreted these findings to indicate that the mothers' monitoring skills in relation to their children's behaviour were deficient, and suggested that interventions that focus partly on improving parents' sensitivity to the full range of their children's behaviour would doubtless be beneficial to such families. These results also suggest that where primary caregivers are inconsistent in their monitoring of their children's behaviour, parental-report methods should not be relied upon in isolation in the assessment of their children's difficulties.

Oldershaw and colleagues observed 20 mothers with their children, with half of the mothers having had a history of physically abusing their children (Oldershaw, Walters & Kordich-Hall, 1986). These researchers found that the mothers who had physically abused their children used 'power-assertive' control (i.e., threats, negative physical, humiliation, disapproval and negative demand) and fewer positive (i.e., reasoning, bargaining, co-operation, modelling and approval) strategies than did those in the comparison group. The former group was also observed to be more intrusive, more inconsistent and less flexible in their use of parenting techniques, and showed more flattened affect than those in the control group. A later study by the same researchers involved the observation of 73 physically abusive mothers and 43 matched controls with their children. The researchers found that the abusive mothers viewed their children in a more negative light than did the control group, although the children did not differ from the control children on any behavioural measure. The researchers were able to subdivide the abusive parents in terms of their parenting style, as emotionally distant, intrusive or hostile on the basis of their observations and subsequent statistical analyses. These investigators found that the behaviour of these mothers' children could also be subdivided, and the forms of behaviour exhibited by the children depended on the parenting style of their mother (Oldershaw, Walters & Kordich-Hall, 1989).

Webster-Stratton (1994) compared two forms of parenting programmes, involving 78 families. Parents were observed at home with their children pre- and post-treatment, and both parenting skills and child behaviour were systematically coded. The researcher found that whether parents completed a basic videotape-based intervention alone or in combination with a more comprehensive parenting

programme, their interactions with their children were observed to have significantly improved over the course of the treatment. More specifically, use of critical statements significantly decreased, and use of praise and reflective statements significantly increased over the treatment time.

Self-Report Assessment of Parenting Skills

Self-report methods of parenting skill assessment, through the use of questionnaires as opposed to interviews, may be considered the antithesis of observation for the many reasons outlined below. Self-report measures may contain open or closed questions. Many of the advantages of using self-report as a method of assessment of parenting skills apply to both types of questions. However, in order to differentiate self-report measures from interview, the following discussion concerns the former type of assessment, which involves closed questions to which only a limited range of responses are available. For example, responses may be sought using a Likert-type scale, or merely with options of 'yes' or 'no.' Measures of frequency of particular behaviours or discipline practices may also be obtained in this way.

Self-report as a method of parenting skills assessment carries with it some advantages that observation does not. Self-report methods are relatively inexpensive, easy and quick to complete and administer, and require no training to use. This means that this method is highly convenient to use. In comparison to the development of coding systems to accompany observations, the development of a self-report measure is on the whole comparatively straightforward. However, as standardised measures of many attractive questionnaires are available, researchers may not need to develop a

questionnaire at all, and this is also increasingly true in the case of observational systems (e.g., Irvine, Biglan, Smolkowski & Ary, 1999; Arnold *et al.*, 1993). In addition, although the behaviours and practices of interest are decided in advance, as is usually true for observational assessment, the participant will have much more control on the information obtained. Further, often minimal or no contact with the investigator is required, it is possible to maintain the participants' anonymity using this method and the associated burden for both the administrators and the participants is relatively small. Hence, this assessment approach may be more appealing to potential participants than research involving observational methods.

It is because of the convenience of self-report use that this method is appropriate for a range of purposes, including screening and treatment evaluation. Although not true of measures specifically designed to assess parenting skills, there are a large number of self-report measures available, such that one can easily employ a combination of self-report measures to assess a wide range of factors at any one time. In contrast to observational methods that may provide finely detailed information about a small range of behaviours over a limited time period, self-report measures allow for a much larger range of behaviours to be assessed across a larger time frame and various situations, and require less time to complete. This, and the fact that many more individuals can be assessed in a give time period by this method compared to by observation, leads to a greater data set than that achieved through observation alone.

Participants can complete the measures without any contact with the investigators, such that this method may be more convenient and thus appealing to

many. It is possible to maintain the participants' anonymity using this method, which may also encourage more parents to participate in the research.

Questionnaires usually elicit a restricted range of answers, such that it is unlikely that researchers will find many of the responses ambiguous. In contrast, observers may well differ in their interpretations of the behaviour of interest, and this is the reason that a high degree of inter-rater reliability and intensive training are necessary in the development and use of a coding system.

However, self-report as an assessment method carries with it a well-documented vulnerability to participant biases, that may lead to an inaccurate set of data. For example, participants may give what they think are model answers to questions rather than answering truthfully. This may relate to the assessment of parenting skills in particular, as parents may be reluctant to admit to certain practices such as physical punishment for which they may expect to be sanctioned given current public feeling about such methods (Arnold *et al.*, 1993). Alternatively, participants may tend to answer all questions in the same way, at the extremes or at the centres of Likert-type scales. There is a range of possible reasons for these response patterns, including low motivation, lack of time or a wish to sabotage the research. Participants may inadvertently give different information to that solicited if any questions appear ambiguous. Other potentially confounding factors include mental health difficulties and external stress (e.g., Webster-Stratton & Spitzer, 1991). In addition, ability to provide accurate self-reports usually depends on the insight, understanding and literacy of the participants, as well as their memory if they are required to provide information retrospectively. Limitations in any of these aspects may well lead to

inaccurate data (e.g., Smith and Brooks-Gunn, 1997; Arnold & O'Leary, 1997; Benjamin, Benjamin & Rind, 1996).

Researchers have made attempts to minimise the likelihood of, or at least detect such response biases. Other methods such as observation or other's report are used by many to support participants' reports. However, researchers not employing multiple methods of assessment may instead make use of special bias test scales, routinely offer assurance to participants that there are no right or wrong answers, and make every effort possible to make the process of providing the information easier for the participants (e.g., Arnold *et al.*, 1993; Eysenck & Eysenck, 1975; Crowne & Marlowe, 1960). Nevertheless, biases in responses or other reasons for potential inaccuracy of the information collected should always be considered when interpreting self-report data, and this remains a serious limitation of this approach (Webster-Stratton & Spitzer, 1991).

The use of closed questions allows researchers to limit the amount of information provided by participants, and this may be advantageous in terms of simplified analysis and data collection process. However, whilst information about a wide range of variables or practices may be collected at one time using this approach, the detail of this data is also more likely to be limited. Where detail of response is limited through the use of closed questions and forced-choice options, self-report measures of assessment may be considered relatively insensitive to detecting small changes in behaviours over time (Taylor *et al.*, 1998).

Finally, low response rates of 30 to 50 percent are normally expected when potential participants are required to return self-report measures via the postal system, and this can make the operation more expensive than it need be (Clark-Carter, 1997).

Examples of Self-Report Based Parenting Research

Benjamin and colleagues (1996) developed the 128-item Likert-type scale Subjective Experiences of Parenting Scale, which aims to assess a range of 14 parenting characteristics (Benjamin *et al.*, 1996). Ninety-four participants completed the questionnaire, 54 of whom experienced dissociative disorders, whilst the remainder formed the control groups. Compared to non-dissociative participants, dissociative individuals reported significantly more negative parenting behaviour and associated parenting attitudes than did the control group. The researchers believed their results lent support to the use of interventions aimed at parenting skills, and suggested that this approach may be especially useful for those who have dissociative disorders.

Arnold and colleagues employed self-report measures to assess paternal parenting and involvement with their children with Attention Deficit Hyperactivity Disorder, or ADHD (Arnold, O'Leary & Edwards, 1997). Among other self-report measures related to the parent-child relationship, these investigators used the Parenting Scale, a 30-item Likert-type scale measure of parenting style, with factors relating to laxness, over-reactivity, verbosity and general dysfunctional discipline (Arnold *et al.*, 1993). Fathers' reports of the time that they spent with their children were also sought. Statistical analyses revealed that fathers reported higher levels of

involvement with their children, the use of more effective discipline, and more love for their wives when the fathers themselves experienced no symptoms of ADHD. This was the first study of its kind in focusing on the relationship between fathers' parenting practices and the amount of time that they spend with their children.

A further example of parenting research using self-report only can be found in the work of Noll and associates, who investigated the child-rearing practices of parents, or caregivers of children with sickle cell disease, or SCD (Noll, McKellop, Vannatta & Kalinyak, 1998). This group used the Child-Rearing Practices Report Q-sort (Block, 1981), a 91-item measure for assessing values, attitudes, and goals of parents about child rearing. The Q-sort employs a card-sorting method of response organisation, but was presented to the participants in the form of a questionnaire with Likert-type scales. The participants also completed other illness assessment measures. Medical professionals involved in the children's care and matched caregiver control individuals (whose children did not have a chronic illness) also took part in the study. These researchers found that, whilst caregiver reports of parenting practices were similar for both groups, professionals perceived the parents of children with SCD to be more protective, more worried and less effective with discipline than did the caregivers themselves. These results were consistent with other similar studies, and researchers interpreted the findings as indicative of a stereotyped view held by medical professionals concerning the negative effects of chronic illness in child-rearing. However, these researchers also highlighted another intriguing possibility, specifically that these caregivers were unaware of their own parenting practices, a finding also generally supported by other research (see above).

MULTIMETHOD PARENTING RESEARCH

Use of observational as well as self-report measures in the assessment of parenting is not uncommon: the growing evidence for their combined validity referred to earlier has no doubt contributed to this trend. However, very few researchers have attempted to compare observational and self-report data regarding the same parenting behaviours, although other methods of assessing the validity of self-report or observational measures have been employed. For those who have attempted such a comparison, this aspect of the study has tended to be secondary to the main purpose of the research, and the results are mixed (Gardner, 1998). Indeed, although Patterson included self-report questionnaire data as part of his pioneering study, he later removed it as this data did not correlate at all well with observational and other measures (Patterson, 1992).

Direct Comparisons of Observational and Self-Report Parenting Data

In developing the Parenting Scale, Arnold and colleagues videotaped 15 mothers with their children at home, eight of whom were in the clinical group (mothers who reported extreme difficulties in coping with their children). All of the parents were asked to interact in a brief free-play period, with the use of toys provided by the researcher. The mothers were then presented with three tasks: to engage in a block-sorting task with their children, throughout which the mothers were instructed to prevent their children from playing with the researcher's toys; persuading the children to clear away the toys; and making a telephone call for 10 minutes. Following the observation of the mothers with their children, the observers rated the

mothers' behaviour in a way that was comparable with the Parenting Scale factors (i.e., according to the degree of laxness, over-reactivity and verbosity demonstrated by the mothers). The non-parametric correlations between the observational and corresponding self-report data were significant, and ranged from 0.53 (Verbosity) to 0.65 (Over-reactivity). However, the researchers acknowledged the need for a replication of this aspect of the study due to the small number of participants.

Kochanska and colleagues compared a community sample of 68 mothers in terms of their self-reported values, attitudes and goals in relation to child-rearing with observations of their interactions with their children in an apartment acquired for the study (Kochanska, Kuczynski & Radke-Yarrow, 1989). The self-report measure employed was the Q-Sort which, as previously outlined, allows participants to group their responses. This measure also seeks responses according to Likert-type scales for each question, and it can also be presented as a questionnaire (Noll *et al.*, 1998). The observers rated both the mothers' strategies, including among others direct commands, reprimands, bargains and physical enforcements, and the children's responses. As mentioned earlier, the Q-sort is an assessment of parental values, attitudes, and goals in relation to child-rearing. These researchers found that the child-rearing attitudes endorsed by the mothers using the Q-sort were positively correlated to their actual child-management strategies. More specifically, mothers whose attitudes were authoritarian or restrictive were more likely to be observed using direct and restrictive strategies. Further, self-reported democratic mothers used relatively indirect, positive and non-confrontational forms of control during the observation phase, but not physical enforcements, prohibitive interventions or direct commands (Kochanska, Kuczynski & Radke-Yarrow, 1989). These results may

indicate that the parents' reports were supported by the observations, although the picture would have been much clearer had the observational and self-report variables of interest been more closely matched.

Smith and Brooks-Gunn (1997) compared the reports of non-clinical mothers regarding their use of physical punishment in disciplining their children with observers' ratings of the same behaviours during home visits. The mothers and children were seen at home six times over the three years of the children's life to the time of the study, with the last visit having taken place at the time the self-report data was collected. Seven-hundred and fifteen mothers and observers were merely asked to report whether the mother had used such discipline methods over a given period. These researchers found that the 293 mothers reporting the use of physical punishment represented a much lower number than expected based on a previous survey. Further, 57 mothers were observed to exert harsh physical discipline, which was a much smaller number than those reporting such practices. However, the correlations between observer and self-reports of use of harsh discipline were reportedly high, as approximately 75 percent of the mothers observed hitting their children during the home visits also reported using more than one physical punishment over the previous week.

Observational and Self-Report Parenting Data Collected but not Compared

Some researchers have gathered both observational and self-report data regarding parenting behaviour, but did not report any comparisons of the two sets. For example, in 1998, Webster-Stratton conducted comprehensive evaluations of a parenting skills programme involving 394 mothers (294 of whom completed parenting skills programme, the remainder represented the control group). All of the mothers responded to questionnaire items concerning three styles of discipline: harsh (including use of verbal or physical aggression or prolonged confinement); consistent (including items relating to consistency in follow-through and predictability of parental responses); or positive (use of verbal encouragement or reinforcement for the child's positive behaviours). Each item was accompanied by a Likert-type scale. The mothers were also systematically observed at home with their children. Mothers attending a parenting programme were observed to show significant decreases in critical and harsh parenting, and increases in positive and competent discipline whilst the control mothers did not, and these findings were supported by similar trends in the mothers' reports of their discipline style. Unfortunately, however, the two measures of discipline style were not compared, such that it is not possible to examine the relationship between the two sets of data.

Arnold and O'Leary (1997) researched maternal and paternal discipline of their children, whom the parents found difficult to manage. These investigators assessed the discipline strategies of 38 parents individually, through self-report and observation of each parent with their child. The self-report measure employed was the Parenting Scale, described previously. Observers rated the parents' behaviour in a

way that reflected two of the subscales of the Parenting Scale (over-reactivity and laxness). Both observational and self-report measures indicated that the mothers were significantly more over-reactive than were the fathers, and the degree to which the mothers as a group were over-reactive was deemed problematic (Arnold & O'Leary, 1997). No difference in laxness across gender was found, and again this was true for both self-report and observational data. However, the two sets of data were not compared directly.

Greenberger and colleagues assessed the parenting behaviour of 188 parents via self-report and observation, during a study concerning the nature of the parents' employment in relation to parenting behaviours (Greenberger, O'Neil & Nagel, 1994). The 39-item purpose-developed questionnaire assessed parental control, with sub-categories of flexible, harsh and lax control. Each item was accompanied by a Likert-type scale. Laboratory observations were directed at the assessment of parents' warmth, responsiveness, and quality of explanations in relation to their children. The researchers found that positive aspects of the parents' employment were generally significantly related to both reported and observed developmentally sound parenting approaches. However, the latter two sets of data were not compared to one another, and indeed such comparisons may have been somewhat irrelevant to the present research given that the different assessment methods were focused on different aspects of the parenting role.

CONCLUSIONS

Externalising behaviours in children are common, and when enduring are difficult to treat and are costly to society. Parenting is thought to play an important role in the development and maintenance of such difficulties. Interventions aimed at parenting appear promising, although research into this area is ongoing. Research involving parental discipline style has often been conducted using observational methods, which have many advantages, but this approach also has disadvantages, and self-report measures have emerged in an attempt to overcome some of these.

There is a clear lack of research involving the comparison of observational and self-report measures of the same aspects of parenting, yet such research is necessary in a comprehensive assessment of the validity of self-report measures. Existing research findings indicate that parental self-report and observational data concerning the same parenting behaviours are not at odds with one another, but that the agreement between them is not complete.

THE PRESENT STUDY

The present study sought to investigate further the concurrent validity of a popular self-report measure of parental discipline style, by examining the relationship between parents' reports of these aspects of their parenting behaviour and objective observations of the same behaviour. The results of such a comparison should have implications in terms of the appropriateness of the use of self-report data in the

assessment of parenting behaviour, and also in relation to parents' ability to report honestly and accurately with respect to parenting behaviours.

The present study was conducted within the context of a broader study, which aimed to investigate the effectiveness of a parenting skills programme. In the larger study, parents and their children were assessed in the manner described in the next section before and after they participated in such a programme. The data employed in the present report represents part of the information gathered at the first (pre-programme) assessment.

The existing research findings suggest that there is some form of relationship between self-report and observational measures of parenting behaviour. However, this association does not appear consistent enough to enable one to predict with any confidence that parental reports of their use of discipline would be significantly and positively related to independent observations of the same behaviours in the present research. Hence, the nature of the investigation into this relationship was exploratory, as described in the next section.

METHOD

Participants

The group of participants comprised children between two and nine years, and one parent, identified as the main caregiver. The children experienced significant behavioural difficulties, as identified by a score of eleven or more on the Problem Scale of the Eyberg Child Behaviour Inventory (Robinson, Eyberg & Ross, 1978: see below), as recommended by Eyberg and Ross (1978).

All of the parents were waiting to start one of a number of parenting programmes run by a charitable organisation, having been referred to the organisation either by themselves or by health or social services personnel for help with parenting skills. All of the parents approached about the research had been assured of a place on either an immediate parenting skills programme, or one to be run in six months' time. The participants were recruited by facilitators of the parenting skills programmes, as a routine part of their assessment.

Design

A within-groups correlational design was used, where observed and reported variables at one time point were compared.

Measures

Both self-report and observational methods were employed, as outlined below.

Self-Report Measures

EYBERG CHILD BEHAVIOUR INVENTORY (ECBI – Robinson *et al.*, 1978)

The 36-item ECBI is a well-known and often used questionnaire, and aims to measure a range of commonly reported children's behavioural difficulties, including defiance, temper tantrums, attention difficulties and bullying of others (*vid.* Appendix 1). These difficulties are assessed on two dimensions: the frequency of the behaviour and the extent to which it is viewed as a problem. The behaviour frequency, or 'Intensity,' is assessed via a seven-point Likert-type scale, ranging from 'never occurs' to 'always occurs.' The total score ranges from 36 to 262. Whether the child's behaviour is viewed as problematic is assessed through a dichotomous 'yes'/'no' response, with a total score range from zero to 36. On both scales, higher scores indicate more behavioural problems.

This measure was consistently found to have good reliability, as demonstrated via test-retest, split half and parallel forms of reliability assessment. The validity of this measure has also been demonstrated in terms of its ability to discriminate clearly between children with and without significant behavioural difficulties. Further, the ECBI has been found to have good construct, criterion-related and internal validity (Robinson, Eyberg & Ross, 1980; Eyberg & Ross, 1978).

THE PARENTING SCALE (Arnold *et al.*, 1993)

This 30-item, three-scale questionnaire was developed as a measure of parental discipline practices that relate theoretically to the development and maintenance of children's behavioural difficulties, as outlined in the previous section (*vid.* Appendix 1). As well as providing a general measure of parental discipline style, the Parenting Scale purports to measure laxness, over-reactivity, and verbosity. Laxness is defined as the ways in which parents give in, allow rules to go unenforced, or provide positive consequences for misbehaviour. An example of a question in this scale is 'When I want my child to stop doing something, I firmly tell my child to stop / I coax or beg my child to stop.' Over-reactivity relates to parental behaviour such as displays of anger, meanness, and irritability. For example, one of the questions of this scale is as follows: 'When my child misbehaves, I raise my voice or yell / I speak to my child calmly.' Verbosity describes the extent to which parents issue lengthy verbal responses and rely on talking even when talking is ineffective (Arnold *et al.*, 1993). An example of a question from the Verbosity scale is: 'If saying no doesn't work right away, I take some other kind of action / I keep talking and try to get through to my child.'

Each item is accompanied by a seven-point, Likert-type scale, used to indicate to what extent the respondent agrees with a range of statements. Several of the items are reversed so that sometimes, the extreme right hand side of the scale gains the highest score per item, and at other times the opposite is true. Substantial overall and scale scores indicate that the parents are making 'discipline mistakes' thought

counterproductive in managing the children's behavioural difficulties (Arnold *et al.*, 1993).

The reliability of the questionnaire was demonstrated via good item-to-factor correlations and impressive test-retest statistics, for a sample of parents whose children were aged between two and four years old. The Laxness, Over-reactivity and overall scores were found to be significantly higher for parents whose children displayed marked behavioural difficulties, compared to the control group, which suggests that it has good construct validity. Further, each of the Parenting Scale total and subscale scores were found to be significantly related to parental reports of levels of child misbehaviour for a sample of 77 parents, whilst only the Over-reactivity scale scores correlated significantly with parental reports of depression (via the Beck Depression Inventory – see below). The Parenting Scale factor scores were found to be highly correlated with objective observations of the same discipline practices (Arnold *et al.*, 1993) although, as discussed in the previous section, the sample involved was small, and the observational coding methods utilised were crude. The present research sought to replicate and extend these findings, with a larger sample with respect to the observational study, and a more subtle observational coding system.

THE PARENTING SENSE OF COMPETENCE SCALE (PSOC – Gibaud-Wallston & Wandersman, 1978)

Based on the available literature, one would expect that the more confidence a parent has in their ability to handle their children's behavioural problems successfully,

the more the same parent would adopt effective parenting practices, and vice versa. Further, one might also expect that the more effectively a parent is able to manage their children's behaviour, the more enjoyment the parent would gain from his or her relationship with the children. Hence, a measure of self-esteem and self-efficacy in relation to parenting, such as the PSOC (*vid.* Appendix 1), could be used as a test of the validity of a discipline style questionnaire such as the Parenting Scale, and this is the reason that the former was employed in the present research.

Originally created in 1978 by Gibaud-Wallston and Wandersman, the PSOC was modified in 1989 following analysis of the measure's psychometric properties (Johnston & Mash, 1989). The modified PSOC is a 16-item questionnaire, with two subscales to measure efficacy and satisfaction in relation to parenting. The PSOC can be found in Appendix 1, although the format of this version differs from the original both in terms of the loss of one item following Johnston and Marsh's work, and also the wording of the items has been altered to make the statements easier to understand. Each item, presented as a statement, is accompanied by a six-point Likert-type scale from 'strongly agree' to 'strongly disagree.' For the self-esteem subscale, the more the respondent agrees with the statements, the higher the score they gain. The opposite is true for the self-efficacy subscale. Higher scores for both subscales indicate higher sense of parenting competency on the part of the respondent.

Johnston and Mash (1989) explored the properties of the 16-item PSOC through the use of a non-clinical sample of community parents, and found that the PSOC had high internal consistency. These researchers also found that the total PSOC scores were moderately but significantly and negatively related to parental

perceptions of their children's behavioural difficulties, which suggests that the PSOC has acceptable construct validity, particularly given the non-clinical nature of the sample.

BECK DEPRESSION INVENTORY (BDI: Beck, Ward, Mendelson, Mock & Erbaugh, 1961)

A review of the relevant literature would lead one to predict that parental depression would affect their ability to discipline their children effectively. Such a relationship may be mediated by a deficit in monitoring skills and responsiveness, low motivation, or a tendency to dwell on the negative aspects of the children's behaviour. Hence, a self-report measure of depression could be used to assess the validity of the a measure of discipline such as the Parenting Scale, and one would expect that an individual who scores highly for depression would also gain high scores on the Parenting Scale. It is for this reason that the BDI, a well-known and commonly used measure of depression, was included in the present study (*vid.* Appendix 1).

Originally developed in 1961, and revised in 1971, the BDI is a 21-item questionnaire, formed on the basis of clinical observations. It was designed to assess the intensity of the depression on the basis of its main symptoms. These include, among other symptoms, sadness, pessimism, guilt, suicidal ideation, and physical symptoms such as loss of appetite and interest in sex. The respondent is asked to chose one of four statements per item that best describes the way he or she has been feeling over the previous week. Respondents gain from zero to three points per item, with a total score range of 21 to 63. A score of between 10 and 18 is considered to

indicate mild to moderate depression, a score of between 19 and 29 is thought to suggest the presence of moderate to severe depression, and a score of higher than 30 is considered to be indicative of severe depression.

Many researchers have assessed the psychometric properties of the BDI, and high internal consistency has been repeatedly established for this measure. The BDI has also been found to have good criterion validity, and moderate to high concurrent and convergent validity. The BDI has also been found to have impressive discriminant validity, at least in terms of its ability to discriminate depressed people from non-depressives (Richter, Werner, Heerlein, Kraus & Sauer, 1998).

Observational Measure

As noted in the previous section, it appears important when assessing the concurrent validity of self-report measures of parenting to use observational measures that are closely linked to the self-report factors under investigation. Otherwise, direct comparisons of self-report and observational data are likely to be of limited value when testing a tool's validity. Whilst there are in existence coding schemes that are applicable to observation of parental discipline strategies, none appeared entirely appropriate. Hence, a purpose-designed coding system was developed for the current research. In addition, it was believed that a more sensitive measure of observed behaviour than that used in the original 1993 validation of the Parenting Scale by Arnold and colleagues could lead to more detailed interpretations of any results.

The present author worked with a research team to develop a detailed coding scheme for the purpose of assessing the concurrent validity of the Parenting Scale. First, the parental behaviours of interest were identified through close examination of the Parenting Scale items. All attempts were made to maximise the match between observed behaviours and those assessed by the Parenting Scale: the observational behaviours are listed in Table 3 along with the Parenting Scale items to which they correspond. For practical reasons, it was not possible to observe all forms of behaviour assessed through the Parenting Scale. Hence, low frequency behaviours (such as the parent apologising to the child when dealing with a problem) or those difficult to rate (for example, whether the parent knows what the child is doing when he or she is out of sight) were not chosen as those to code for analysis. However, all other parental discipline behaviours thought to be observable within the time available were included within the coding system, and all three of the Parenting Scale subscales (i.e., Laxness, Over-reactivity and Verbosity) were represented. These behaviours included: the issue of all forms of command; use of threats and the extent to which these were followed through; incidences and outcome of conflict; any parental utterances following a command whilst still in conflict with the child; raised voice or negative tone; coaxing or begging, and rudeness directed at the child; and aggression. These behaviours are described in detail in Appendix 2. The coding grid used during the observations can be found in Appendix 3.

Once appropriate observable behaviours were identified, the team met regularly over the course of ten months with the aims of reaching consensus on the definition of the behaviours to be observed, and to achieve acceptable inter-rater

Observed Behaviour [*]	Parenting Scale Item	Parenting Scale Item Content	Parenting Scale Subscale
Following threats through	7	Not following threats through	Laxness
Commands (all types)	8	Limit setting	Laxness
Coaxing / begging	12	Firmness of command	Laxness
Conflict outcome	19	Not keeping to limits set	Laxness
Following threats through	20	Not following threats through	Laxness
Conflict outcome	26	Not keeping to limits set	Laxness
Conflict outcome	30	Not keeping to limits set	Laxness
<u>verbosity</u>	6	Arguing with the child	Over-reactivity
<u>verbosity</u>	9	Use of lectures	Over-reactivity
Negative tone / raised voice	10	Raised voice	Over-reactivity
Aggression	18	Use of physical punishment	Over-reactivity
Rudeness	25	Use of 'bad language'	Over-reactivity
Rudeness	28	Use of insulting terms aimed at child	Over-reactivity
<u>verbosity</u> / vague commands	4	Amount said when telling child not to do something	Verbosity
<u>verbosity</u>	7	Realistic/unrealistic threats	Verbosity
<u>verbosity</u> / vague commands	9	Amount said when child misbehaves	Verbosity
<u>verbosity</u>	29	Verbal response to complaints from child	Verbosity

Table 3: Observed (Coded) Behaviours with Corresponding Parenting Scale Items

reliability for each variable. These aims were worked towards through practice coding of videotaped interactions and detailed discussion of any variability in the

^{*} see Appendix 2 for full descriptions of behaviours

coders interpretations, with reference to the available literature in defining some forms of behaviour or commands (e.g., Webster-Stratton, 1994). Thus, the team trained together in the behavioural coding whilst the coding system was in the process of development.

In this way, the observational coding scheme for the present research was formed. In order to make the coding process more manageable, the behaviours of interest were coded for every 30-second period. However, only total frequency counts (i.e., for the entire period of observation) were used in the main analyses.

It was important that attempts to discipline the children were differentiated from non-optimal play. For this reason, coding was only undertaken during the structured tasks (video, joint play and skittles - see below) if conflict unrelated to the play took place at these times. When this happened, all coding resumed for the duration of the conflict episode. Hence, coding of observed behaviours took place during transitions (i.e., at the beginning and end of each structured task), periods of conflict, and at all times outside structured tasks.

The reliability of the one individual who undertook the coding of all the videotapes (the author) was assessed through comparing this coder's ratings with that of another member of the research group who had also been greatly involved in the development of the coding scheme. The outcome of this reliability analysis is reported in the Results section.

Procedure

Participants were given detailed information about the research by those assessing them for parenting programmes. A member of the research team then contacted those who were interested in participating and, if the potential participant was still willing to take part in the research, a home visit was arranged. Participants were required to give written consent before the data collection began. A copy of the consent form can be found in Appendix 4.

A member of the research team would then visit the family at home. All of the required materials were provided by the team member. Although the focus of the observation, and data collection in general, was on the main caregiver and the index child, the presence of other family members during the research assessment was not discouraged in any way. The research visit comprised several stages. Following a short period during which the format for the visit was outlined, and the family members were able to become accustomed to the researcher's presence, a series of short tasks began. The tasks were carefully chosen to maximise the opportunity to view the behaviours of interest, which were expected throughout the tasks but at transition points in particular.

First, the index child was asked to watch a popular form of cartoon, one of a series of animations. After approximately five minutes, when the first cartoon had finished, the parent was instructed to tell the child to turn the television and video off. Next, the parent was asked to engage in ten minutes of joint play with the child and any other children present if desired, using a set of farm toys provided by the

researcher. When this period had drawn to a close, the parent was asked to engage the index child in clearing up the toys. Once the toys had been cleared away, the parent and child were given some skittles, and they were instructed that they could decide who, along with the child, would be involved in the skittles game. After five minutes, the parent was asked to tell the child to clear the skittles away. The final stage of the research visit was a period of 20 minutes, during which the parent was asked to complete the questionnaires and then do whatever they would normally do were the researcher not present. The entire visit was recorded by videotape for later coding, with the participants' explicit consent.

Ethical Issues

There were ethical issues associated with this research that required addressing. Firstly, there is a possibility that parents would consent to the research because they believed that this decision would influence the likelihood of attending a parenting skills programme in the near future. However, all potential participants were assured that they would be invited to attend a programme within the next six months, whether participating in the research or not. Whether they would attend a programme immediately or in six months' time was dependent on the random allocation, and the participants were made aware of this.

A second issue was the possibility that participants may have been concerned that others could identify them as having children with behavioural difficulties. To prevent this, all the information collected, whether via self-report or through

observation, was kept securely at the research base, and numbers rather than names were attached to the questionnaires and the videotapes.

Thirdly, there is a possibility that participants may have become distressed through the completion of questionnaires or through the observation part of data collection. The questionnaires elicited personal information concerning the relationship between the parents and their children, and the parents may have found the process difficult, particularly where their relationship with their children was strained. Equally, the tasks that the parents and children were expected to undertake, as described above, would potentially lead to conflict between parent and child: indeed, the tasks were selected in the hope that some mild conflict may arise during the observation. However, the tasks were deemed likely to be pleasant for, and perhaps familiar to many families, and it is probable that the transitions between tasks during the research visit were representative of those that occurred multiple times everyday. Nonetheless, the participants were made aware of the need for their continued consent from the initial stage, such that if they wanted to withdraw from the research at any point, they were free to do so.

As mentioned previously, the present research was secondary to research involving the evaluation of a parenting programme, and did not involve any additional data collection or contact with the participants. Ethical consent for the wider study was sought from, and granted by, the Applied and Qualitative Research Ethics Committee (Oxfordshire Health Authority: *vid.* Appendix 5).

The results of the statistical analyses are outlined in the following section.

RESULTS

Description of Sample

Data from 40 families who participated in the larger study were analysed. Power analyses indicated that this number of participants was sufficient for high-powered (>0.8) correlational analyses which would be able to detect an effect size of 0.3 or higher.

Parents

The sample of parents involved in this study may be summarised according to key features, as shown in Table 4.

Age and Gender	
Age Range	20 – 49 years (mean =30 years, 1 month)
Gender Ratio (M:F)	1:19
Marital Status	
Marital Status	Married = 15, Divorced = 9, Separated = 3
Number with a Partner	29
Number of Partners living with the family	19
Employment	
Number in Full-Time Employment	1
Number in Part-Time Employment	15
Number Studying	1
Number Unemployed	23
State Benefits	
Number in Receipt of State Benefits	18
Social Class	
Number Representing Social Classes 2 and 3 (minor professional, managerial, non-manual)	16
Number Representing Social Classes 3 – 5 (manual, semi-skilled and unskilled)	13

Table 4: Description of Parents whose Data was Analysed in the Study (N=40)

As can be seen in Table 4, 12 of the sample were lone parents, and most were not employed outside the home. Almost half of the 40 parents were claiming state benefit (14% is currently the national average for claiming such benefits). A little over half of the parents represented minor professional, managerial and non-manual social classes, whilst 13 of the parents represented manual, semi-skilled and unskilled social classes. Taken together, these statistics indicate that this was a very disadvantaged sample.

Children

General characteristics of the index children involved in this study are summarised by Table 5. As Table 5 shows, 27 male and 13 female index children participated in the study. These children were aged between two years, three months and eight years, five months. Whilst 28 of the children did have siblings, 12 of the index children had none.

Age and Gender	
Age Range	2 - 8 years (mean = 5 years, 2 months)
Gender Ratio (M:F)	27:13
Position in Family	
Number of Lone Children	12
Number of First (Eldest) Children	16
Number of Siblings	
Number of Siblings (Range)	0 – 5 (mean = 2)

Table 5: Description of Index Children (N=40)

Behaviour Coding System Reliability Analysis

The development of the behaviour coding system employed as part of the present study involved a lengthy training period with the aim of limiting inter-observer variability, as mentioned previously. However, it is possible that despite such training, observational data sets, such as that involved in the current study, do not accurately reflect the participants' behaviour, due to coder bias or an under-developed coding system. In order to explore this possibility, it is common practice to assess the level of agreement, or inter-rater reliability, between two or more observers of the same events. Indeed, such analysis is considered essential in observational research (Margolin, Oliver, Gordis, O'Hearn, Medina, Ghosh & Morland, 1998; Shrout & Fleiss, 1979). A high level of agreement between coders would indicate that they are consistently identifying the occurrence, or non-occurrence, of the same behaviour, and this suggests that the observational data set does accurately reflect the behaviour of the participants (e.g., Hops, Davis & Longoria, 1995; Repp, Deitz, Boles, Deitz & Repp, 1976).

Choice of Method

There are several methods of assessing inter-rater agreement, and choice of method depends greatly on researcher preference. However, one may be guided by the data set itself, as it appears important to match the method of agreement analysis to the way in which the data set is organised for the final analyses. Hence, whilst sequential analyses require stringent point-by-point agreement, overall agreement on

total frequency of behaviour is acceptable for non-sequential analyses, that is, analyses involving only overall frequency count (Margolin *et al.*, 1998).

Methods of calculating inter-rater agreement that are appropriate for non-sequential overall frequency count analysis include percentage agreement, Pearson's correlations, and intraclass correlations. Percentage agreement involves dividing the number of observations that the coders agreed upon by the total number of observations made, and presenting this as a percentage. Good inter-rater agreement is implied through high percentage agreement values. However, this method does not control for chance levels of agreement between observers, such that high percentage agreement does not necessarily mean that the coding system employed is a reliable one.

Similarly, Pearson's product-moment correlations can be used to indicate the extent to which two coders are reliable. This involves a calculation of the degree to which the coded frequencies of behaviour co-vary across observers. The associated correlation coefficients range from -1.00 to $+1.00$. A correlation coefficient approaching or equal to $+1.00$ is an indication of good inter-rater reliability. Unfortunately, this method is insensitive to differences across observers in the levels of the coded behaviour, such that a high correlation coefficient may be gained when observers are consistent in their differences, both in terms of size and direction. This means that one can not assume from a high correlation coefficient that the observers are even coding the same behaviours.

The intraclass correlation procedure is also often employed to assess the agreement between raters where total frequency data is involved. This method compares the difference in observers' ratings of behaviour to differences in the participants in terms of the measured behaviour. A high intraclass correlation indicates that the variance due to coders is low relative to that due to differences among participants on the observed behaviour, and that either observer's data distinguishes equally between participants over the given time period (Margolin *et al.*, 1998; Bakeman & Gottman, 1997). As this popular method was appropriate for the observational data set of the present study, and was not vulnerable to positive findings occurring by chance or due to consistent disagreement as other methods were (e.g., percentage agreement or simple correlations), it was employed in this research.

Preparations for the Inter-Rater Reliability Analysis

For the reliability analysis, the main researcher and another member of the research team rated the behaviours of 15 participants (37.5% of the present sample) for 10 minutes per participant, according to the behavioural coding system outlined previously. Power analyses indicated that this number of participants was sufficient for high-powered (>0.75) analyses-of-variance type procedure, which would be able to detect an effect size of 0.2 or higher. The same participants and identical segments of the videotaped home visit were observed and coded by the two researchers, and were randomly chosen. The two researchers' total frequencies of the behaviour for the 10-minute periods formed the data set for the reliability analysis.

Key characteristics of this data set, comprising the observational variables for the reliability analysis, are displayed in Appendix 6. The distribution of the observational data set was first examined. Some frequencies of particular behaviours were combined where appropriate (see Table 3), and this compensated for low-frequency data for individual variables. Follow through of threat was not examined due to absence of such behaviour in the coded segments of the home visits. Measures of skewness and kurtosis for the observations were assessed at a 0.01 level of significance, which allows for some degree of skewness and kurtosis common in data collected via observational methods. Nonetheless, most of the distributions were significantly skewed and had kurtosis. Transforming the data made little difference in this respect. The behavioural frequencies were then combined to represent the Parenting Scale discipline style factors they were designed to reflect. Laxness was a function of begging/coaxing, total number of commands, and outcome of conflict, such that higher frequencies of begging/coaxing, fewer commands, and fewer occasions where parents 'won' a conflict represented a greater degree of Laxness. Over-reactivity comprised verbosity (i.e., utterances following commands: indicated by a small, underlined initial y in this report to distinguish it from the group variable, Verbosity. See Appendix 2 for full descriptions of observed behaviours), yelling, rudeness and aggression, such that parents were found to be more over-reactive when they were verbose, shouted at their children, and were rude or aggressive towards their children. Verbosity combined the vague commands index and verbosity. Square roots were taken for all groups.

These transformed data sets were considered statistically normal, according to the ratio of skewness or kurtosis index to standard error ($p < 0.01$). However, the

results of the Kolmogorov-Smirnov test, which tests the null hypothesis that the data is not normally distributed, indicated that whilst transformed Laxness and Verbosity data were normal ($p=0.20$), one of the Over-reactivity data sets was not normally distributed. Further transformations of this data set made little difference in this respect, according to the Kolmogorv-Smirnov test of normality.

It is argued that the use of intraclass correlations in assessing inter-rater reliability regarding the observational variables is most appropriate for variables meeting the assumption of normality. However, the data set for this part of the analysis was small, and much of it was inevitably skewed and had kurtosis. As there is no non-parametric alternative to the intraclass correlation test of inter-rater reliability, and as such a parametric analysis of variance is robust and so tolerant of a certain degree of non-normality, intraclass correlations were performed both for the subgroups outlined above and for the individual behaviours. However, it is acknowledged that, where the non-normal variables were subject to reliability analysis, the power of the test may have been compromised.

Reliability Analysis

The two-way, mixed effect reliability analysis then took place. The basic bivariate and intraclass correlations between observers for the main discipline style subgroups and also for individual behaviours (as outlined above) are displayed in Table 6.

Observed Discipline Style/Variable	Bivariate Correlation Coefficient	Intraclass Correlation Coefficient
Laxness*	0.986	0.983
Over-reactivity*	0.886	0.830
Verbosity (Vague Commands plus verbosity)*	0.927	0.931
Total Number of Commands	0.985	0.967
Child Compliance	0.871	0.874
Child Non-compliance	0.990	0.972
Non Vague Commands	0.867	0.870
Yelling (Yell + Yell+)	0.958	0.950
Rudeness + Aggression	0.888	0.885
Begging/Coaxing	0.969	0.922
Conflict	0.942	0.943
Conflict Outcome ('parent wins conflict' minus 'child wins conflict')	0.834	0.845

* The square root of the (positive) discipline style subgroup was used in the analysis, as this improved the normality of the associated distribution

Table 6: Bivariate and Intraclass Reliability Coefficients for Observed Discipline Styles and Variables

As shown in Table 6, both the bivariate and the intraclass correlations for the discipline style subgroups and other variables were extremely high, indicating that the two raters were sufficiently reliable in their observations. Hence, it appears that the observational data set does accurately represent the behaviour of the participants. The full observational data set could thus be used in the main analysis, where the concurrent validity of the Parenting Scale was assessed.

Analysis of the Concurrent Validity of the Parenting Scale

Preparations for the Concurrent Validity Analysis

Following the positive results of the reliability analysis as reported above, the complete observational and questionnaire data sets for the 40 families were then assessed for normality in preparation for the main analyses, as recommended by Clark-Carter (1997). Scrutiny of relevant histograms and analysis of the normal probability plots for each questionnaire factor indicated that, with the exception of the satisfaction scale of the PSOC, all factors were normally distributed. When the reciprocal of the PSOC satisfaction scale was obtained, this associated distribution was also normal. The observational discipline style subgroups (i.e., those representing Laxness, Over-reactivity and Verbosity) were not normally distributed, but when square roots of these variables were taken, the associated distributions were found to be normal. Use of the Kolmogorov- Smirnov test conformed the suitability of all of the main variables for correlational analyses, once the above transformations had taken place ($0.103 < p < 0.200$). The mean, minimum and maximum values, in addition to the standard deviation for each key variable are displayed in Table 7.

Correlational Analysis: Parental Discipline Style

The observed and the parent-reported discipline style frequencies were then subject to correlational analysis.

Origin of Variable	Variable	Mean	Minimum	Maximum	Standard Deviation
Parental Behaviour	Laxness ϕ *	-76.95	-272	-14.00	56.89
Parental Behaviour	Over-reactivity*	10.37	0	30.00	8.66
Parental Behaviour	Verbosity*	44.33	1.00	170.00	37.53
Parenting Scale (PS)	Laxness	33.14	15.00	64.00	11.85
PS	Over-reactivity	36.91	13.00	57.00	10.20
PS	Verbosity	27.49	11.00	43.00	6.99
Parenting Sense of Competence (PSOC) scale	Satisfaction \diamond	27.68	17.00	40.00	5.98
PSOC	Efficacy	25.74	14.00	38.00	6.19
PSOC	Total Score	53.74	38.00	73.00	10.39
Beck Depression Inventory	Total Score	19.14	0	39.00	11.60
Eyberg Child Behaviour Inventory (ECBI)	Intensity	162.46	99.00	209.00	24.56
ECBI	Problem	19.60	10.00	32.00	5.76

ϕ The negative values associated with this variable are due to the way in which the composite was constructed (see above); * The square root of the (positive) variable was used in the analysis; \diamond The reciprocal of the variable was used in the main analysis

Table 7: Summary Statistics for Key Variables in Main Analysis

Whilst the observed discipline styles were highly and significantly correlated with one another ($0.74 < r < 0.96$), these relationships were partly due to shared behaviours or items across subgroups. The parent-reported discipline styles were also interrelated but to a lesser extent ($0.36 < r < 0.50$), for the same reason of shared items.

As a test of the reliability of the observational coding system, further non-parametric correlational analyses were conducted to assess the extent to which the observed discipline styles were related to the individual behaviours of which they were formed. One would expect the observational subgroups to correlate highly with

the individual behaviours of which they were made up, if the behavioural subgroupings were appropriate. Only the number of commands made a significant contribution to the Laxness discipline style: the fewer commands made, the higher degree of Laxness was present ($r = -0.993$, $p < 0.001$). Neither begging/coaxing nor conflict outcome was related to Laxness. Over-reactivity was significantly related to three of its four composites, in the expected direction, verbosity ($r = 0.88$, $p < 0.001$), yelling ($r = 0.85$, $p < 0.001$) and rudeness ($r = 0.64$, $p < 0.001$). Aggression was not entered into the analysis due to low frequency of occurrence (three times in total, for two parents). Finally, Verbosity was also significantly related to its constituent behaviours in the expected direction, verbosity ($r = 0.98$, $p < 0.001$) and vague commands ($r = 0.83$, $p < 0.001$).

The results of the two-tailed parametric correlational analysis for the observed and the corresponding parent-reported discipline styles are displayed in Table 8. Although the relationship between observed and reported over-reactivity approached significance ($r = 0.29$, $p = 0.07$), none of the observed discipline styles were significantly related to any of those from the Parenting Scale.

		PARENTING SCALE DISCIPLINE STYLE		
		Laxness	Over-reactivity	Verbosity
OBSERVED DISCIPLINE STYLE	Laxness*	-0.13 (NS)	0.06 (NS)	-0.03 (NS)
	Over-reactivity*	-0.06 (NS)	0.29 (NS)	0.16 (NS)
	Verbosity*	-0.12 (NS)	0.09 (NS)	0.00 (NS)

* The square root of the (positive) variable was used in the analysis; NS = non-significant

Table 8: Results of Correlational Analysis of Parental Discipline Style (Pearson)

Correlational Analysis: Observed and Parent-Reported Discipline Behaviours

The concurrent validity of the Parenting Scale was investigated further, through correlational analysis involving individual Parenting Scale items and their corresponding observed behaviours. The aim of this additional analysis was to establish the degree to which the Parenting Scale items were statistically related to the observed behaviours to which they corresponded theoretically.

As stated previously, not all of the observational variables were appropriate for parametric analyses, due to large degrees of skewness and kurtosis. Hence, both parametric and non-parametric correlational analyses were employed where appropriate in the analyses. The distributions of both observed variables and data relating to Parenting Scale items were examined using the methods outlined earlier, and data were transformed where necessary (square roots were taken in all cases). The frequencies of the variables were also taken into account, such that variables of low frequency (follow through of threat, and aggression) were withheld from the analysis, as they were considered unsuitable. Indeed, a threat was observed to be followed through only once, for only one parent, and aggression towards an index child was only observed three times, for two parents. Summary statistics for the relevant behavioural and Parenting Scale variables can be found in Appendix 7.

The behavioural variables were then correlated with the Parenting Scale items to which they corresponded. The results of this set of two-tail parametric and non-parametric analyses are summarised by Table 9.

Observed Behaviour	Parenting Scale Item	Content of Parenting Scale Item	Correlation Coefficient	Parametric (P) or Non-parametric (N) Analysis
Verbosity ϕ (vague commands + verbosity)	4	Amount said when telling child not to do something	-0.07 (NS)	N
	9	Amount said when child misbehaves	-0.02 (NS)	N
verbosity	6	Arguing with the child	0.34*	P
	9	Amount said when child misbehaves	0.05 (NS)	N
	29	Verbal response to complaints from child	-0.02 (NS)	N
Yelling (Yell + Yell+)	10	Raised voice	0.20 (NS)	N
Rudeness	25	Use of 'bad language'	0.12 (NS)	N
	28	Use of insulting terms aimed at child	0.36**	N
Begging/Coaxing	12	Firmness of command	0.10 (NS)	N
Conflict ϕ	6	Arguing with the child	0.06 (NS)	P
Conflict Outcome ϕ (‘parent wins conflict’ minus ‘child wins conflict’)	19	Keeping to limits set	0.06 (NS)	P
	26	Keeping to limits set	0.02 (NS)	N
	30	Keeping to limits set	-0.02 (NS)	N

ϕ The square root of this variable was used in the analysis; * $p < 0.05$, ** $p < 0.01$; NS = non-significant

Table 9: Results of Correlational Analysis of Observed and Parent-Reported Discipline Behaviours

As Table 9 summarises, most of the observed discipline behaviours were not significantly related to the Parenting Scale items to which they corresponded. The exceptions to this were verbosity, which was related to reported engagement in arguments with the child, and Rudeness, which was correlated with reported use of insulting terms aimed at the child.

Correlational Analysis: Parenting Scale and Other Factors

The Parenting Scale factors and other parent-report factors were then subject to two-tailed, parametric correlation analysis, the results of which are displayed in Table 10.

		PARENTING SCALE DISCIPLINE STYLE		
		Laxness	Over-reactivity	Verbosity
OTHER PARENT-REPORTED FACTORS	PSOC Satisfaction	-0.35*	-0.49**	-0.27 (NS)
	PSOC Efficacy ♦	0.36*	0.47**	0.24 (NS)
	PSOC Total Score	-0.38*	-0.63**	-0.12 (NS)
	BDI Total Score	0.35*	0.28 (NS)	0.00 (NS)
	ECBI Intensity	0.35*	0.18 (NS)	0.18 (NS)
	ECBI Problem	0.28 (NS)	0.10 (NS)	0.13 (NS)

* $p < 0.05$, ** $P < 0.01$; NS = non-significant; ♦ The reciprocal of the variable was used in the main analysis, such that a positive correlation with this variable indicates a negative relationship

Table 10: Results of Correlational Analysis of Parenting Scale and Other Factors (Pearson)

As can be seen in Table 10, Parenting Scale Laxness and Over-reactivity were significantly and negatively correlated with parental sense of satisfaction, self-efficacy in relation to parenting, and general parenting competence. Laxness was the only Parenting Scale factor to be significantly related to parent-reported depression and parent-perceived child behaviour problem intensity, and these relationships were in the positive direction. Parenting Scale Verbosity was not significantly related to any of the other parent-report factors. The extent to which parents viewed their

children’s behaviour as problematic was not correlated with any of the Parenting Scale factors.

Correlational Analysis: Observed Parenting and Other Factors

The observed parenting styles and other parent-reported factors were then subject to two-tailed, parametric correlational analyses, the results of which are outlined in Table 11.

		OBSERVED PARENTING STYLE		
		Laxness*	Over-reactivity*	Verbosity*
OTHER PARENT-REPORTED FACTORS	PSOC Satisfaction	0.04 (NS)	0.05 (NS)	0.37 (NS)
	PSOC Efficacy♦	0.06 (NS)	0.07 (NS)	0.12 (NS)
	PSOC Total Score	0.04 (NS)	0.05 (NS)	0.09 (NS)
	BDI Total Score	0.00 (NS)	-0.06 (NS)	-0.02 (NS)
	ECBI Intensity	0.11 (NS)	0.15 (NS)	0.13 (NS)
	ECBI Problem	-0.15 (NS)	0.07 (NS)	-0.11 (NS)

* The square root of the (positive) variable was used in the analysis; ♦ The reciprocal of the variable was used in the main analysis; NS = non-significant

Table 11: Results of Correlational Analysis of Observed Parenting Behaviour Subgroups and Other Parent-Reported Factors (Pearson)

As Table 11 shows, the observed parenting subgroups were not significantly related to parenting sense of competence, parental depression, or level of child behaviour problems perceived by the parent.

All of the above results are discussed in detail in the next section.

DISCUSSION

The present study sought to investigate further the concurrent validity of the Parenting Scale, by examining the relationship between parental self-report and objective observations of discipline style. Comparisons of observational and self-report measures of the same aspects of parental discipline style have rarely been reported, and where such work has been undertaken, results have been mixed.

In the present study, the results of the initial inter-rater reliability analysis were good, as the reliability coefficients were high both for individual behaviours and for those grouped corresponding to Parenting Scale discipline styles (i.e., Laxness, Over-reactivity and Verbosity). This indicates that two observers consistently identified the same behaviour relating to the participants. This suggests that the coding system employed did lead to an observational data set that accurately reflected the behaviour of the participants.

A good test of the internal reliability of an observational coding system involves assessing the degree to which behavioural subgroups (in this case, observed discipline style) represent their constituent parts (the individual observed behaviours). An analysis of this kind was undertaken as part of the present study. The results of this analysis revealed that only the number of commands made a significant contribution to the Laxness observational subgroup, in the expected direction: the more commands made, the less Laxness was present. Neither begging/coaxing nor conflict outcome made a statistically significant contribution to the Laxness subgroup.

Over-reactivity was significantly related to three of its four composites (verbosity, yelling, and rudeness) in the expected direction: higher Over-reactivity frequencies were associated with more observed verbosity, yelling or rudeness. Aggression was not entered into the analysis due to low frequency of occurrence. Finally, Verbosity was also significantly related to its constituent behaviours in the expected direction, as parents who were verbose or who tended to issue vague commands regularly gained a higher overall Verbosity frequency count.

In the main analysis, the observed behavioural discipline styles were not significantly related to the corresponding self-report subscales of the Parenting Scale, although the relationship between the observed Over-reactivity behavioural subgroups and the corresponding parent-reported factors did approach significance. The relationships between individual observed behaviours and their corresponding Parenting Scale items were also examined in order to assess the degree to which the Parenting Scale items were statistically related to the observed behaviours to which they corresponded theoretically. Only two of the individual observed behaviours were significantly correlated with their Parenting Scale counterparts. These behaviours were continuing to talk or argue with the child following the issuing of a command (before the end of the conflict episode), and rudeness (particularly the use of insulting terms) directed at the child.

The relationships between Parenting Scale and other parent-reported factors were also examined. This analysis revealed that parent-reported Laxness and Over-reactivity were significantly and negatively related to measures of parental sense of competence. The more that a parent was lax or over-reactive towards their child, the

less the parent felt satisfied and competent in their parenting, and the less they felt their parenting was efficacious. Laxness alone was related to parental depression and the intensity of perceived child behaviour difficulties, such that the more that a parent was lax in their discipline style, the more depressed they considered themselves to be, and the worse they perceived their children's behavioural difficulties to be. However, the results were mixed, as parent-reported verbosity was not significantly related to any of the other parent-reported factors, and there were no significant relationships between observed parenting discipline style and any of the other parent-reported variables listed above.

The concurrent validity of the Parenting Scale, or the degree to which it accurately reflects the discipline behaviours of its respondents, will be discussed first with reference to the above results. Following this discussion, the validity of the observational system will also be explored. This is because the meaning and implications of the above results depend greatly on the validity of the observational system employed in this study. These two discussions will then lead to interpretation of the results taken together. Finally, clinical implications, and areas for future research will be outlined.

The Concurrent Validity of the Parenting Scale

The coding system was specifically designed to reflect behaviours referred to in the Parenting Scale. However, as outlined above, no significant relationship was found between any of the observed and parent-reported disciplinary styles, and only

two individual observed behaviours were significantly correlated with their Parenting Scale counterparts. These results are in contrast to those reported by Arnold and colleagues in the original Parenting Scale validation study (Arnold *et al.*, 1993). These researchers found that the Parenting Scale factor scores were strongly and significantly related to observers' general ratings of the same behaviours.

The methodology of the present study differed from that of Arnold and colleagues' observational work (1993) in two key respects. First, the sample employed in the present study comprised 40 parents of children with behavioural problems that were clinically significant. In contrast, the observational part of the original validation study involved a sample of 15 mothers with their children, where only eight of the mothers reported extreme difficulties in coping with their children. Second, Arnold and colleagues (1993) employed a much less detailed observational rating system than that used in the present study. Their observational data were based on the observers' global impressions of the parents' behaviour, involving only overall ratings of how lax, over-reactive or verbose the observers felt that the mothers had been in the videotaped interactions with their children. For these reasons, it is suggested that the present study represented a more appropriate and rigorous test of the concurrent validity of the Parenting Scale than did the observational work in the original validation study (Arnold *et al.*, 1993). Hence, the lack of relationships between parents' reports and observations of disciplinary behaviour in the present study suggests that the Parenting Scale does not have good concurrent validity.

Results of comparisons of parents' responses to the Parenting Scale and parent-reported child behaviour, parental sense of competence, and depression lent

some support to this view. Although there were some significant relationships between Parenting Scale and other parent-report variables in a direction one might have expected, these findings were not consistent across all factors. This pattern of results does not fully reflect those gained in the original Parenting Scale validation study, in which only Over-reactivity was significantly associated with parental report of depression, whilst all three subscales were significantly correlated with a parent-report measure of child behaviour difficulty (Arnold *et al.*, 1993). Given the well-established validity of measures such as the Beck Depression Inventory, or the Eyberg Child Behaviour Inventory, it would appear that the contrasting results outlined above are in some way related to the use of the Parenting Scale. This view is supported by the finding that the data gained through the use of this questionnaire was not supported by the objective observations.

In this study, the Verbosity factor of the Parenting Scale was not related to any of the other parent-reported variables. Similarly, Arnold and colleagues (1993) reported that the validity results for the Verbosity factor were mixed, and Verbosity did not distinguish children on the basis of level of behavioural problems. It may be that, compared to other disciplinary styles, verbosity is a weak factor. Perhaps the extent to which verbosity in particular is an ineffective discipline approach depends greatly on other contextual variables, such as the relationship between the parent and child or the severity of the conflict.

Lack of concurrent validity is a serious flaw in any measure, as this suggests that the tool is not measuring the variables that it was designed to assess. However, in order to test the concurrent validity of one measure, the other instrument against

which it is compared must have good validity. As mentioned previously, the observational system used in the present study was a more detailed measure of behaviour than that employed by Arnold and colleagues (1993). However, if the observational coding system employed in this study was not a valid measure of the behaviours of interest (i.e., those highlighted in the Parenting Scale) then one cannot consider comparisons between the two to represent a good test of the concurrent validity of the questionnaire. If this were the case, it would mean that one could not conclude from this study that the Parenting Scale is not sufficiently valid. Further, an observational system with poor validity could explain most of the findings reported above, specifically the lack of relationship between Parenting Scale and observed behaviours, and also the finding that none of the observational factors were significantly related to any of the other parent-reported variables. Hence, the validity of the observational coding system must be examined before any solid conclusions about the Parenting Scale are formed.

The Validity of the Observational Coding System

- **Face validity**

Poor face validity of the observational coding system used in this study (due to Parenting Scale behaviours not being accurately represented by the observed behaviours) would indicate that the employment of the system in this study was not appropriate as a test of the concurrent validity of the Parenting Scale. However, the face validity of the coding system was good. Though not exhaustive, the behaviours to be observed were carefully chosen to reflect those included in the Parenting Scale,

to a similar degree of detail (e.g., whether a parent enters into an argument with the child, rather than the exact number of words used in response to the child). Further, the reliability analysis indicated that the two observers were consistently identifying the same behaviour, such that either set of observations could reliably distinguish one participant from another. The validity of such a coding system is strongly related to its reliability, although impressive reliability is not by itself sufficient to imply that the system is valid in all respects.

- Choice of settings or tasks

The validity of observational coding system could also have been threatened by the choice of settings or tasks. If these were inappropriate, the coding system may not have been registering behaviour that would occur naturally. However, the authors of the Parenting Scale used similar tasks in their validation study, so one may expect to gain similar results in the present study. Further, it is suggested that whilst observations of parents and children during structured tasks cannot automatically be generalised to more natural circumstances, it is nonetheless likely that the parents would attempt to manage their children's behaviour using their usual methods, regardless of the task. In addition, observer presence does not appear to alter parents' behaviour significantly (Gardner, 2000; see the Introduction to the present study). Hence, it is suggested that neither the chosen tasks nor the fact that the parents were under observation would significantly influence the parents' behaviour in this study.

- Limited data set

Observational methods do tend to lead to small behavioural data sets particularly in comparison to the data gained through the use of questionnaires (Gardner, 2000), and this applies to the present study. It is noteworthy that some of the individual observed behaviours occurred infrequently. In particular, following through on a threat occurred rarely, often because the children responded to the threat by complying with the original command, such that following through on the threat was not necessary. The analysis of the internal reliability of the coding system revealed that most of the individual behaviour frequencies did make significant contributions to the relevant overall discipline style scores. However, parental aggression, begging or coaxing, or conflict outcome (defined as the number of times the parent 'won' in conflict episodes minus the number of times that the child 'won') were not significantly related to the discipline styles that were thought to represent them (i.e., Over-reactivity and Laxness). These results may indicate that these behaviours or positive parental conflict outcome occurred too infrequently to make a difference to the observed discipline style frequencies. It has already been acknowledged that parental follow through of threat, and aggression occurred infrequently during the observation period. Consideration of the frequencies of begging or coaxing and positive parental conflict outcome (see Appendix 7) suggests that these did occur for a significant proportion of the parents observed, although they did not occur often within single observation periods. It may be that an extended observation period would have led to higher frequencies of observed behaviour: the length of the observation period is discussed in more detail below.

- Inappropriate groupings of individual behaviours as a measure of discipline style

There is an alternative interpretation of results of the analysis of the internal reliability of the coding system, outlined above. These findings may indicate that the inclusion of parental aggression, begging or coaxing, or conflict outcome in the observed discipline style subgroups was not appropriate, because the individual variables do not relate theoretically to the discipline subgroups which were considered to represent them. Since the observed behaviours were chosen, defined and grouped according to Parenting Scale factors, if the above assertion is true of the observational system, it must also be true of the Parenting Scale. However, good internal reliability of the Parenting Scale factors has been demonstrated using a large sample of respondents (Arnold *et al.*, 1993). Hence, it appears that the groupings of behaviours for both the Parenting Scale and the observational system used in this study did reflect the discipline styles accurately.

- Observational period

It may be that, though the coding system itself may have been generally valid, the observation period (approximately one-hour) was insufficient, leading to limited observational data sets, as mentioned above. This is a difficulty related to the use of observation methods in general, as discussed in the Introduction to the present research: however long or detailed an objective observation period may be, information can only be gained about a small part of the lives of those being observed. However, one hour of observation is comparable to that involved using other coding systems, and was recommended by Patterson (1982). Further, other observational

coding systems involving similar lengths of observation period have been shown to have good validity, and are sensitive to changes in parental behaviour following intervention. In addition, the structured tasks were employed in order to stimulate mild conflict, such that it was more likely that the behaviours of interest would occur during the observation period. Hence, it appears that the observation period was sufficient for most of the behaviours of interest, although there still remains a difficulty in objective assessment of less frequent behaviours.

Interpretation of the Results

Since it seems likely that the observational system was sufficiently valid and appropriate for this study, the results of this research suggest that the Parenting Scale itself indeed lacks concurrent validity. Possible explanations for this are discussed below.

Problems Inherent in the Design of the Parenting Scale

It may be that the individual behaviours described by the Parenting Scale do not accurately reflect the discipline styles that this questionnaire purports to measure (i.e., Laxness, Over-reactivity and Verbosity), or that the content-validity of the Parenting Scale is also questionable due to a limited range of items per factor. Another possibility is that discipline styles other than laxness, over-reactivity and verbosity are more relevant in the development of child behaviour difficulties, parental sense of competence and parental depression. The finding that the

corresponding observed variables were not correlated with any of the other parent-reported variables supports this second view.

The Effect of Respondent-Related Factors on the Validity of the Parenting Scale

Another possibility is that the contrasting results gained in the present study and the initial validation research are due to differences in the samples involved in the two investigations. Arnold and colleagues (1993) found that their clinical sample did not differ significantly from their non-clinical sample in terms of demographic details. In contrast, the sample involved in the current study was clearly markedly disadvantaged in terms of the high proportion of lone parents, the number of parents claiming state benefits, and the distribution of the sample in terms of social class. Although the perceived level of children's difficulties were comparable in both studies, the participants in the study of Arnold and colleagues (1993) viewed themselves as markedly less depressed than did the those in the present study. Further, despite finding key differences in their clinical and non-clinical groups in terms of these same variables, Arnold and associates (1993) combined both groups in their analysis of concurrent validity. This regrouping served to make Arnold and colleagues' (1993) sample more heterogeneous, and this further differentiates their sample from the one used in the present study.

The age ranges for the children for each study were also different (aged two to four years for Arnold and colleagues, 1993; aged two to eight years for the present study). However, it seems somewhat unlikely that the difference in children's ages between the two studies should account wholly for the differential results, because the

age ranges of the children in these studies overlapped. In addition, there is a lack of theoretical reason to suppose that parents would differ enormously in their discipline styles for children between two and eight years, though they may adapt their technique somewhat according to the age of the child.

Deficits in Parental Monitoring Skills

The lack of relationship between observed and reported disciplinary behaviours may be explained by a lack of parental sensitivity, or monitoring skill deficit. As mentioned previously, parental difficulties in monitoring their children's behaviour have been linked with children's behavioural problems (e.g., Carr, 1999; Webster-Stratton & Herbert, 1994; Wahler and Sansbury, 1990). Some researchers have argued that parenting skills are likely to be limited for parents who lack the ability to monitor, and therefore respond to, their children's moment-to-moment behaviour accurately (Wahler & Dumas, 1989). Difficulties in monitoring children's behaviour may be related to narrowed attention directed at the child, coupled with a tendency to attend to perceived child misbehaviour rather than more positive behaviour on the child's part. These factors would lead to a lack of recognition of the complexity of the situation, resulting in a tendency to form simplistic judgements about the reason that the child is behaving as he or she is (e.g., 'he is just a naughty child'). This in turn would shape and limit the discipline strategies employed by the parent (Wahler & Dumas, 1989). At the same time, it may be that these parents also experience difficulties in monitoring their own behaviour, particularly where they have employed disciplinary approaches inconsistently. Inconsistency in parental disciplinary behaviours has also been associated with the development of children's

behavioural problems (e.g., Carr, 1999). A lack in ability to monitor behaviour would doubtless influence the degree to which parental reports accurately reflected their disciplinary practices, and may also explain why self-reported discipline behaviours were not related to other reported variables such as depression or parenting sense of competence.

An immediate question raised by this hypothesis is what would make some parents more vulnerable to narrowed or biased attention, as well as a tendency to jump to unfounded conclusions about their child's, and also their own, behaviour. As discussed in the Introduction to this research, many factors are believed to be associated with the development of children's behavioural difficulties via their influence on parenting style. These include parental mental health, the quality of the parents' relationship with one another, beliefs about parenting methods and abilities, parental attribution style, external stress, and available support. As discussed above, the parents involved in the present study were markedly depressed and socially disadvantaged. It is possible that all of these and other stresses work in combination with the initial development of the child's behavioural difficulties to overwhelm the vulnerable parent, leading to demoralisation, a reduction in motivation and perceived coping resources and abilities, and biases in their attention, thinking and memory. It seems highly possible that these changes would affect the parent's ability to attend to the child's behaviour and needs, and may well render them less able to monitor their own behaviour towards their children (e.g., Webster-Stratton & Herbert, 1994; Wahler & Dumas, 1989).

Biases in Response to Parenting Scale

The possible memory bias that may influence the responses to retrospective questionnaires such as the Parenting Scale is well documented. It is possible that parents were not accurately remembering how they behaved towards their children, or were concentrating only on the previous few weeks, where their behaviour may not have been representative of their normal disciplinary practices for some reason. In addition, it may have been that parents' responses to the Parenting Scale items were affected by social desirability bias: perhaps parents tended to describe their disciplinary behaviour in a way they thought was acceptable. Indeed, Arnold and colleagues raised this issue themselves in their original validation study report (Arnold *et al.*, 1993). However, the wording and order of the Parenting Scale items was carefully planned to prevent any 'right' answers from being too obvious (Arnold *et al.*, 1993). Further, given that the parents involved in the present study had been referred for help with parenting skills, it seems likely that their beliefs regarding 'good' discipline techniques were wide ranging. The range of responses to the Parenting Scale items (see Appendix 7) supports this view.

Clinical and Research Implications

The results of the present research suggest that the Parenting Scale does not consistently provide an accurate measure of parental discipline practices via parental self-report. Whatever the explanation for these results, it seems that one should be cautious when employing the Parenting Scale in future studies, and the use of an additional measure of discipline style (e.g., objective observation or reports of other

family members) appears necessary. Further investigation of the Parenting Scale, and perhaps other self-report measures of parental discipline style, is recommended. Indeed, until such work is undertaken, results gained through the use of any self-report measures of discipline style in parenting research should be subject to cautious interpretation. At the same time, the Parenting Scale may be used informally, in highlighting areas of concern for parents, promoting discussion about different discipline styles, and perhaps aiding clinicians and parents alike in identifying an initial focus for intervention.

Another intriguing interpretation of the above results is that the parental reports of discipline on which the present study focussed were not valid, perhaps due to a deficit in monitoring skills, which may in turn be due to factors such as parental depression or demoralisation. There appears to be some support for this view from research into other aspects of family behaviour. For example, Achenbach and colleagues found in a meta-analysis that there was only a small overall degree of association ($r=0.27$) between parent report of their children's emotional and behavioural difficulties and objective observations of the same difficulties (Achenbach, McConaughy & Howell, 1987). Another research group found that parents significantly overestimated the amount of time that their children spent watching television, as determined by comparing parental reports with videotaped observations (Anderson, Field, Collins, Lorch & Nathan, 1985). However, parents' ability to monitor their own parenting practices has received little attention from researchers to date, yet this appears to be an important area for further investigation. Results of such research may have clear implications for treatment aimed at the

development of parenting skills, as well as for research studies involving the assessment of the effectiveness of such interventions.

Parenting interventions with a focus on raising parental awareness of their own behaviours as well as those of their children seems appropriate. Indeed, many parenting programmes achieve this heightened sensitivity to behaviour secondary to the development of parenting skills (e.g., Webster-Stratton, 1992). Further research involving the use of the Parenting Scale before and after such a programme may lead to positive changes in the accuracy of the parental reports of discipline. Where such programmes are not readily available, parents may benefit from concentrating on enhancing their monitoring skills. There may be a role for clinicians here, as they could initially work with parents of children with behavioural difficulties to increase their ability to monitor behaviour through the use of diaries or frequency charts, for example. Clinicians can also work with such parents in developing stress management skills, which may in turn augment their monitoring abilities (Wahler & Dumas, 1989). Finally, as acknowledged earlier, interventions directed at the development of parenting skills should be offered as part of a wider treatment package, aimed at addressing other important factors such as other family relationships and social isolation.

Limitations

The inclusion of another self-report measure of discipline style may have been illuminating, as a further test of the concurrent validity of the Parenting Scale,

although the different measures would claim to assess different aspects of discipline style.

Although the coding system employed appears to have been a sufficiently valid observational measure of parental discipline style as defined by the Parenting Scale, it is likely that the system is not a valid measure of parental discipline in general. This is because the validity of the Parenting Scale and its factors that the coding system was carefully designed to test has been called into question. Further, contrary to what one may have expected, none of the observational measures were correlated with other well-established parental report measures of other variables. Also, this coding system was quite limited as, due to practical reasons, only a restricted range of behaviours could be observed (see the Method section). At present, the coding system may be used in observations of a specific range of disciplinary behaviours, and further validation studies may indicate that it is a good instrument to be used for this purpose. For example, the wider research with which the present study was associated will involve the use of the coding system pre- and post- intervention, to assess whether it is sensitive to changes in behaviour with treatment. However, it appears that further research into, and development of, the coding system described above is needed before it is employed again as an assessment of particular parenting styles, rather than a measure of individual discipline behaviours.

The sample size appeared adequate for the purpose of assessing the validity of the Parenting Scale. However, a larger sample would have helped to increase the power of the correlational analysis. Further, many of the distributions associated with

the individual behaviours were not considered normal, and neither were they easily transformed. Although again the size of the sample involved in the reliability analysis seemed appropriate for the statistical procedure employed, a larger sample would have increased the power of the reliability analyses, and may have led to distributions more appropriate for parametric analyses.

Conclusions

Although the Parenting Scale is a popular measure of parenting discipline style, it appears that the most likely explanation of the results outlined above and in the previous section is that this instrument does not have sufficient concurrent validity for general use. This finding contrasts with the results of the original Parenting Scale validation study (Arnold *et al.*, 1993). The observational aspect of the original study involved far fewer participants who were characteristically different from those involved in the present research, and a much more basic observational rating system than did the present study. There may be several reasons for the above finding. One intriguing possibility is that some parents find it difficult to monitor their own discipline practices, and this affects their ability to report accurately on how they discipline their children. These results have implications for therapeutic intervention aimed at parenting skills, in terms of early work aimed at enhancing the monitoring ability and stress management skills of parents whose children have behavioural problems. These findings have also led to suggestions for future research involving parental self-report measures of discipline, the Parenting Scale in particular. Further, this study has highlighted potential areas of further research. These include exploring the validity of other parent-report measures of discipline style and investigating further the role that parental monitoring skills deficits play in the development and maintenance of children's behavioural difficulties. It is hoped that such work will lead to improvements in research and interventions aimed at reducing these difficulties for children and their families.

REFERENCES

- Achenbach, T., McConaughy, S. & Howell, C. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213-232
- American Psychiatric Association (APA) (1994). *Diagnostic and Statistical Manual (Version 4) (DSM-IV)*. Washington, D.C.: APA
- Anderson, D., Field, D., Collins, P., Lorch, E. & Nathan, J. (1985). Estimates of young children's time with television: a methodological comparison of parent reports with time-lapse video home observation. *Child Development*, *56*, 1345-1357
- Arnold, E. & O'Leary, S. (1997). Mothers' and fathers' discipline of hard-to-manage toddlers. *Child and Family Behavior Therapy*, *19*(3), 1-11
- Arnold, E., O'Leary, S. & Edwards, G. (1997). Father involvement and self-reported parenting of children with Attention Deficit – Hyperactivity Disorder. *Journal of Consulting and Clinical Psychology*, *65*(2), 337-342
- Arnold, D., O'Leary, S., Wolff, L. & Acker, M. (1993). The Parenting Scale: a measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, *5*(2), 137-144
- Bakeman, R. & Gottman, J. (1997). *Observing Interaction: an Introduction to Sequential Analysis (2nd Edition)*. New York: Cambridge University Press
- Beck, A., Ward, C., Mendelson, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561-571
- Benjamin, L., Benjamin, R. & Rind, B. (1996). Dissociative mothers' subjective experience of parenting. *Child Abuse and Neglect*, *20*(10), 933-942
- Block, J. (1981). The Child-Rearing Practices Report (CRPR): a set of Q items for the description of parental socialization attitudes and values. Unpublished manuscript, Institute of Human Development, University of California, Berkeley
- Bradley, R. & Corwyn, R. (1999). Parenting. In Balter, L. & Tamis-LeMonda, C. (Eds.), *Child Psychology: a Handbook of Contemporary Issues*. Hove: Psychology Press, pp 339-362
- Campbell, S. (1995). Behavior problems in preschool children: a review of recent research. *Journal of Child Psychology and Psychiatry*, *36*(1), 113-149
- Carr, A. (1999). *The Handbook of Child and Adolescent Clinical Psychology: a Contextual Approach*. London: Routledge

Clark-Carter, D. (1997). *Doing Quantitative Psychological Research: from Design to Report*. East Sussex: Psychology Press

Crowne, D.P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, **24**, 349-354.

Dreyer, B., Mendelsohn, A. & Tamis-LeMonda, C. (1996). Assessing the child's cognitive home environment through parental report: reliability and validity. *Early Development and Parenting*, **5(4)**, 271-287

Eyberg, S. & Ross, A. (1978). Assessment of child behavior problems: the validation of a new inventory. *Journal of Clinical Child Psychology*, **7**, 113-116

Eysenck, H. & Eysenck, S. (1975). *Manual of the Eysenck Personality Scales*. London: Hodder & Stoughton

Gardner, F., Sonouga-Burke, E. & Sayal, K. (1999). Parents anticipating misbehaviour: an observational study of strategies parents use to prevent conflict with behaviour problem children. *Journal of Child Psychology and Psychiatry*, **40(8)**, 1185-1196

Gardner, F. (2000). Methodological issues in the direct observation of parent-child interaction: do observational findings reflect the natural behaviour of participants? *Clinical Child and Family Psychology Review*, **3(3)**, 185-198

Gardner, F. (1998). Observational studies of parent-child interaction and behaviour problems: their implications for parenting interventions. In Buchanan, A & Hudson, B. (Eds.), *Parenting, Schooling, and Children's Behaviour*. UK: Ashgate, pp 53-72

Gardner, F. (1997). Observational methods for recording parent-child interaction: how generalisable are the findings? *Child Psychology and Psychiatry Review*, **2(2)**, 70-74

Gardner, F. (1994). The quality of joint activity between mothers and their children with behaviour problems. *Journal of Child Psychology and Psychiatry*, **35(5)**, 935-948

Gardner, F. (1989). Inconsistent parenting: is there evidence for a link with children's conduct problems? *Journal of Abnormal Child Psychology*, **17(2)**, 223-233

Gardner, F. (1987). Positive interaction between mothers and conduct-problem children: is there training for harmony as well as fighting? *Journal of Abnormal Child Psychology*, **15(2)**, 283-293

Gibaud-Wallston, J. & Wandersman, L. (1978). *Development and utility of the Parenting Sense of Competence scale*. Paper presented at the meeting of the American Psychological Association, Toronto

Greenberger, E., O'Neil, R. & Nagel, S. (1994). Linking workplace and homeplace: relations between the nature of adults' work and their parenting behaviors. *Developmental Psychology*, **30**(6), 990-1002

Hops, H., Davis, B. & Longoria, N. (1995). Methodological issues in direct observation: illustrations with the Living in Familial Environments (LIFE) coding system. *Journal of Clinical Child Psychology*, **24**(2), 193-203

Irvine, A., Biglan, A., Smolkowski, K. & Ary, D. (1999). The value of the Parenting Scale for measuring the discipline practices of parents of middle school children. *Behaviour Research and Therapy*, **37**, 127-142

Johnston, C. & Mash, E. (1989). A measure of parenting satisfaction and efficacy. *Journal of Clinical Child Psychology*, **18**(2), 167-175

Kochanska, G., Kuczynski, L. & Radke-Yarrow, M. (1989). Correspondence between mothers' self-reported and observed child-rearing practices. *Child Development*, **60**, 56-63

Margolin, G., Oliver, P., Gordis, E., O'Hearn, H., Medina, A., Ghosh, C. & Morland, L. (1998). The nuts and bolts of behavioral observation of marital and family interaction. *Clinical Child and Family Psychology Review*, **1**(4), 195-213

Noll, R., McKellop, J., Vannatta, K. & Kalinyak, K. (1998). Child-rearing practices of primary caregivers of children with Sickle Cell Disease: the perspective of professionals and caregivers. *Journal of Pediatric Psychology*, **23**(2), 131-140

Oldershaw, L., Walters, G. & Kordich-Hall, D. (1989). A behavioral approach to the classification of different types of physically abusive mothers. *Merrill-Palmer Quarterly*, **35**(3), 255-279

Oldershaw, L., Walters, G. & Kordich-Hall, D. (1986). Control strategies and noncompliance in abusive mother-child dyads: an observational study. *Child Development*, **57**, 722-732

O'Leary, S. (1995). Parental discipline mistakes. *Current Directions in Psychological Science*, **4**(1), 11-13

Patterson, G. (1982). *Coercive Family Process*. USA: Castalia

Patterson, G. & Bank, L. (1986). Bootstrapping your way in the nomological thicket. *Behavioral Assessment*, **8**, 49-73

Reid, J. & Patterson, G. (1989). The development of antisocial behaviour patterns in childhood and adolescence. *European Journal of Personality*, **3**, 107-119

- Repp, A., Deitz, D., Boles, S., Deitz, S. & Repp, C. (1976). Differences among common methods for calculating interobserver agreement. *Journal of Applied Behavior Analysis*, *9*, 109-113
- Richter, P., Werner, J., Heerlein, A., Kraus, A. & Sauer, H. (1998). On the validity of the Beck Depression Inventory: a review. *Psychopathology*, *31*, 160-168
- Robins, L. (1991). Conduct disorder. *Journal of Child Psychology and Psychiatry*, *32(1)*, 193-212
- Robinson, E. & Eyberg, S. (1978). The dyadic parent-child interaction coding system: standardization and validation. *Journal of Consulting and Clinical Psychology*, *49(2)*, 245-250
- Robinson, E., Eyberg, S. & Ross, A. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child Psychology*, *9(1)*, 22-27
- Serketich, W. & Dumas, J. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: a meta-analysis. *Behavior Therapy*, *27*, 171-186
- Shrout, P. & Fleiss, J. (1979). Intraclass reliability: uses in assessing rater reliability. *Psychological Bulletin*, *86(2)*, 420-428
- Simons, R., Whitbeck, L., Conger, R. & Chyi-In, W. (1991). Intergenerational transmission of harsh parenting. *Developmental Psychology*, *27(1)*, 159-171
- Smith, J. & Brooks-Gunn, J. (1997). Correlates and consequences of harsh discipline for young children. *Archives of Pediatric and Adolescent Medicine*, *151*, 777-786
- Smith, A. & O'Leary, S. (1995). Attributions and arousal as predictors of maternal discipline. *Cognitive Therapy and Research*, *19(4)*, 459-471
- Taylor, T., Schmidt, F., Pepler, D. & Hodgkins, C. (1998). A comparison of eclectic treatment with Webster-Stratton's Parents and Children Series in a children's mental health center: a randomized controlled trial. *Behavior Therapy*, *29*, 221-240
- Vuchinich, S., Bank, L. & Patterson, G. (1992). Parenting, peers, and the stability of antisocial behavior in preadolescent boys. *Developmental Psychology*, *28 (3)*, 510-521
- Wahler, R. & Sansbury, L. (1990). The monitoring skills of troubled mothers: their problems in defining child deviance. *Journal of Abnormal Child Psychology*, *18(5)*, 577-589
- Wahler, R. & Dumas, J. (1989). Attentional problems in dysfunctional mother-child interactions: an interbehavioral model. *Psychological Bulletin*, *105(1)*, 116-130

Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, *66*(5), 715-730

Webster-Stratton, C. (1994). Advancing videotape parent training: a comparison study. *Journal of Consulting and Clinical psychology*, *62*(3), 583-593

Webster-Stratton, C. (1992). *The Incredible Years*. Toronto: Umbrella Press

Webster-Stratton, C. (1984). Randomized trial of two parent training programs for families with conduct-disordered children. *Journal of Consulting and Clinical Psychology*, *52*(4), 666-678

Webster-Stratton, C. (1982). Long-term effects of a videotape modeling parent education program: comparison of immediate and 1-year follow-up results. *Behavior Therapy*, *13*, 712-714

Webster-Stratton, C. (1981). Videotape modeling: a method of parent education. *Journal of Clinical Child Psychology*, *10*, 93-98

Webster-Stratton, C. & Herbert, M. (1994). *Troubled Families: Problem Children*. Chichester: Wiley

Webster-Stratton, C., Hollinsworth, T. & Kolpacoff, M. (1989). The long-term effectiveness and clinical significance of three cost-effective training programs for families with conduct-problem children. *Journal of Consulting and Clinical Psychology*, *57*(4), 550-553

Webster-Stratton, C. & Spitzer, A. (1991). Development, reliability, and validity of the Daily Telephone Discipline interview. *Behavioral Assessment*, *13*, 221-239

APPENDICES

Appendix 1: Questionnaires

Appendix 2: Coding Rules

Appendix 3: Coding Grid

Appendix 4: Consent Form

Appendix 5: Letter Indicating Ethical Approval

Appendix 6: Key Characteristics of Variables for Reliability Analysis

Appendix 7: Key Characteristics of Variables for Correlational Analysis Involving Observed and Parent-Reported Parenting Variables

APPENDIX 1

Questionnaires

The Eyberg Child Behaviour Inventory

The Parenting Scale

The Parenting Sense of Competence Scale

The Beck Depression Inventory

Subject No.

Date

Child's name:

EYBERG CHILD BEHAVIOUR INVENTORY

DIRECTIONS:

Below are statements that describe children's behaviour.

- Please:
- (1) Circle the number describing how often the behaviour currently occurs with your child;
 - (2) Circle 'yes' or 'no' to indicate whether the behaviour is currently a problem for you.

	Never	Seldom		Sometimes		Often		Always	Is this a problem for you?
1. Dawdles in getting dressed	1	2	3	4	5	6	7		yes/no
2. Dawdles or lingers at mealtimes	1	2	3	4	5	6	7		yes/no
3. Has poor table manners	1	2	3	4	5	6	7		yes/no
4. Refuses to eat food presented	1	2	3	4	5	6	7		yes/no
5. Refuses to help around the house when asked	1	2	3	4	5	6	7		yes/no
6. Slow in getting ready for bed	1	2	3	4	5	6	7		yes/no
7. Refuses to go to bed on time	1	2	3	4	5	6	7		yes/no
8. Does not obey the house rules on his own	1	2	3	4	5	6	7		yes/no
9. Refuses to obey until threatened with punishment	1	2	3	4	5	6	7		yes/no
10. Acts defiant when told to do something	1	2	3	4	5	6	7		yes/no
11. Argues with parents about rules	1	2	3	4	5	6	7		yes/no
12. Gets angry when doesn't get his own way	1	2	3	4	5	6	7		yes/no
13. Has temper tantrums	1	2	3	4	5	6	7		yes/no

	<div> <div>Never</div> <div>Seldom</div> <div>Sometimes</div> <div>Often</div> <div>Always</div> </div>							Is this a problem for you?
14. Cheeky to adults	1	2	3	4	5	6	7	yes/no
15. Whines	1	2	3	4	5	6	7	yes/no
16. Cries easily	1	2	3	4	5	6	7	yes/no
17. Shouts or screams	1	2	3	4	5	6	7	yes/no
18. Hits parents	1	2	3	4	5	6	7	yes/no
19. Destroys toys and other objects	1	2	3	4	5	6	7	yes/no
20. Is careless with toys and other objects	1	2	3	4	5	6	7	yes/no
21. Steals	1	2	3	4	5	6	7	yes/no
22. Lies	1	2	3	4	5	6	7	yes/no
23. Teases or provokes other children	1	2	3	4	5	6	7	yes/no
24. Argues with friends his own age	1	2	3	4	5	6	7	yes/no
25. Argues with brothers and sisters	1	2	3	4	5	6	7	yes/no
26. Fights with friends his own age	1	2	3	4	5	6	7	yes/no
27. Fights with sisters and brothers	1	2	3	4	5	6	7	yes/no
28. Constantly seeks attention	1	2	3	4	5	6	7	yes/no
29. Interrupts	1	2	3	4	5	6	7	yes/no
30. Is easily distracted	1	2	3	4	5	6	7	yes/no
31. Has short attention span	1	2	3	4	5	6	7	yes/no
32. Fails to finish tasks or projects	1	2	3	4	5	6	7	yes/no
33. Has difficulty entertaining himself alone	1	2	3	4	5	6	7	yes/no
34. Has difficulty concentrating on one thing	1	2	3	4	5	6	7	yes/no
35. Is over-active or restless	1	2	3	4	5	6	7	yes/no
36. Wets the bed	1	2	3	4	5	6	7	yes/no

Subject No.

Date

PARENTING SCALE

Parents have many different ways or styles of dealing with children's misbehaviour (e.g. having a tantrum, whining, hitting someone, answering back, forgetting homework). Below are items that describe some styles of parenting.

For each item cross the box that best describes your style of parenting during the past two months.

SAMPLE ITEM

At meal times

I let my child decide how much to eat

☐☐☐☐☐☐☐☐

I decide how much my child eats

1. When my child misbehaves

I do something right away

☐☐☐☐☐☐☐☐

I do something about it later

2. Before I do something about a problem

I give my child several reminders or warnings

☐☐☐☐☐☐☐☐

I use only one reminder or warning

3. When I'm upset or under stress

I am picky and on my child's back

☐☐☐☐☐☐☐☐

I am no more picky than usual

4. When I tell my child not to do something

I say very little

☐☐☐☐☐☐☐☐

I say a lot

5. When my child pesters me

I can ignore the pestering

☐☐☐☐☐☐☐☐

I can't ignore the pestering









6. When my child misbehaves

I usually get into a long argument with my child

☐☐☐☐☐☐☐☐

I don't get into an argument

- | | | |
|--|---|--|
| <p>7. I threaten to do things that
I am sure I can carry out</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>I know I won't actually do</p> |
| <p>8. I am the kind of parent that
sets limits on what my child
is allowed to do</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>lets my child do whatever he
or she wants</p> |
| <p>9. When my child misbehaves
I give my child a long lecture</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>I keep my talks short and to the
point</p> |
| <p>10. When my child misbehaves
I raise my voice or yell</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>I speak to my child calmly</p> |
| <p>11. If saying no doesn't work right
away
I take some other kind of action</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>I keep talking and try to get
through to my child</p> |
| <p>12. When I want my child to stop
doing something
I firmly tell my child to stop</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>I coax or beg my child to stop</p> |
| <p>13. When my child is out of sight
I often don't know what my
child is doing</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>I always have a good idea of
what my child is doing</p> |
| <p>14. After there's been a problem
with my child
I often hold a grudge</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>things get back to normal quickly</p> |
| <p>15. When we're not at home
I handle my child the way I do
at home</p> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <p>I let my child get away with
a lot more</p> |

- 16. When my child does something I don't like**
- I do something about it every time it happens  I often let it go
- 17. When there's a problem with my child**
- things build up and I do things I don't mean to do  things don't get out of hand
- 18. When my child misbehaves, I spank, slap, grab, or hit my child**
- never or rarely  most of the time
- 19. When my child doesn't do what I ask**
- I often let it go or end up doing it myself  I take some other action
- 20. When I give a fair threat or warning**
- I often don't carry it out  I always do what I said
- 21. If saying no doesn't work**
- I take some other kind of action  I offer my child something nice so he/she will behave
- 22. When my child misbehaves**
- I handle it without getting upset  I get so frustrated or angry that my child can see I'm upset
- 23. When my child misbehaves**
- I make my child tell me why he/she did it  I say "no" or take some other action

24. If my child misbehaves and then acts sorry

I handle the problem like I usually would

--	--	--	--	--	--	--

I let it go that time

25. When my child misbehaves

I rarely use bad language or curse

--	--	--	--	--	--	--

I almost always use bad language

26. When I say my child can't do something

I let my child do it anyway

--	--	--	--	--	--	--

I stick to what I said

27. When I have to handle a problem

I tell my child I'm sorry about it

--	--	--	--	--	--	--

I don't say I'm sorry

28. When my child does something I don't like, I insult my child, say mean things, or call my child names

never or rarely

--	--	--	--	--	--	--

most of the time

29. If my child talks back or complains when I handle a problem

I ignore the complaining and stick to what I said

--	--	--	--	--	--	--

I give my child a talk about not complaining

30. If my child gets upset when I say "No"

I back down and give in to my child

--	--	--	--	--	--	--

I stick to what I said

Subject No. Date **PSOC****Being a parent**

Listed below are a number of statements. Please circle the number beside each statement which most accurately shows how you feel about each statement.

	Strongly agree	Agree	Mildly agree	Mildly disagree	Disagree	Strongly disagree
1. Even though being a parent could be rewarding I am frustrated while my child is at his/her present age	1	2	3	4	5	6
2. I go to bed and wake up feeling that I have not achieved very much.	1	2	3	4	5	6
3. Sometimes when I'm supposed to be in control, I feel more like the one being manipulated.	1	2	3	4	5	6
4. My own mother was better prepared to be a good parent than I am.	1	2	3	4	5	6
5. I would make a good role model for new parents who needed to learn what it takes to be a good parent.	1	2	3	4	5	6
6. Being a parent is manageable, and any problems are easily solved.	1	2	3	4	5	6
7. A difficult thing about being a parent is not knowing whether you are doing a good job or a bad one.	1	2	3	4	5	6
8. Sometimes I feel like I'm not getting anything done.	1	2	3	4	5	6
9. I am satisfied with the way I care for my child.	1	2	3	4	5	6
10. If anyone can find the answer to what is troubling my child, I can.	1	2	3	4	5	6

	Strongly agree	Agree	Mildly agree	Mildly disagree	Disagree	Strongly disagree
11. My talents and interests are in other areas, not in being a parent.	1	2	3	4	5	6
12. Considering how long I've been a parent, I feel completely at home with this role.	1	2	3	4	5	6
13. If being a parent to a child were more interesting, then I would be motivated to do a better job as a parent.	1	2	3	4	5	6
14. I honestly believe that I have all the skills necessary to be a good parent to my child.	1	2	3	4	5	6
15. Being a parent makes me tense and anxious.	1	2	3	4	5	6
16. I now realise that the problems of taking care of a child are easy to solve once you know how your actions affect your child.	1	2	3	4	5	6

Subject No.

Date

BECK INVENTORY

On this questionnaire are groups of statements. Please read each group of statements carefully then pick out the one statement in each group which best describes the way you have been feeling the past week including today. Be sure to read all the statements in each group before making your choice.

NOW READ EACH GROUP OF STATEMENTS CAREFULLY AND MARK THE STATEMENT THAT BEST DESCRIBES THE WAY YOU HAVE BEEN FEELING.

1. ☐ 0 I do not feel sad
☐ 1 I feel sad
☐ 2 I am sad all the time and I can't snap out of it
☐ 3 I am so sad or unhappy that I can't stand it

2. ☐ 0 I am not particularly discouraged about the future
☐ 1 I feel discouraged about the future
☐ 2 I feel I have nothing to look forward to
☐ 3 I feel that the future is hopeless and that things cannot improve

3. ☐ 0 I do not feel like a failure
☐ 1 I feel I have failed more than the average person
☐ 2 As I look back on my life, all I can see is a lot of failures
☐ 3 I feel I am a complete failure as a person

4. ☐ 0 I get as much satisfaction out of things as I used to
☐ 1 I don't enjoy things the way I used to
☐ 2 I don't get real satisfaction out of anything anymore
☐ 3 I am dissatisfied or bored with everything

5. ☐ 0 I don't feel particularly guilty
☐ 1 I feel guilty a good part of the time
☐ 2 I feel quite guilty most of the time
☐ 3 I feel guilty all of the time

6. ☐ 0 I don't feel I am being punished
☐ 1 I feel I may be punished
☐ 2 I expect to be punished
☐ 3 I feel I am being punished

7. ☐ 0 I don't feel disappointed in myself
☐ 1 I am disappointed in myself
☐ 2 I am disgusted with myself
☐ 3 I hate myself

18. ☐ 0 My appetite is not worse than usual
☐ 1 My appetite is not as good as it used to be
☐ 2 My appetite is much worse now
☐ 3 I have no appetite at all anymore
19. ☐ 0 I haven't lost much weight, if any lately
☐ 1 I have lost more than 5 pounds
☐ 2 I have lost more than 10 pounds
☐ 3 I have lost more than 15 pounds
20. I am purposely trying to lose weight by eating less
☐ 0 No
☐ 1 Yes
21. ☐ 0 I am no more worried about my health than usual
☐ 1 I am worried about physical problems such as aches and pains, an upset stomach or constipation
☐ 2 I am very worried about physical problems and it's hard to think of much else
☐ 3 I am so worried about physical problems than I can't think about anything else
22. ☐ 0 I have not noticed any recent change in my interest in sex
☐ 1 I am less interested in sex than I used to be
☐ 2 I am much less interested in sex now
☐ 3 I have lost interest in sex completely

APPENDIX 2

Coding Rules

GENERAL CODING RULES

Begin coding at start of video unless there are more than 5 minutes before start of activities.

Note on 30 second intervals:

A segment runs from 01.00 to 01.29.59 inclusive. Coder should pause the tape at the end of each 30-second segment.

CONFLICT (CF):

Dispute between parent and child. It can begin with:

- a parent or child command or prohibition, plus a non-comply to that
- with a child Annoy behaviour preceded or followed by a prohibition or telling off by parent

When CF is initiated by child (i.e., child demand which is then followed by a prohibition from parent), rate prohibition and match it with a child non-comply.

Conflict consists of a string, lasting for 2 or more turns, e.g., P: "Get me the book," C: "No," of commands, non-complies, or other strictly defined Annoy/Aversive behaviours.

Conflict ends when parent and child are no longer in opposition about something, and the contrary, annoying or organising strings have ended.

If there are only 2 turns of conflict and these span two intervals, this is not coded as CF.

When conflict begins with a prohibition, e.g., P: "Don't climb on the chair" followed by child continuing to climb, this is CF. Command followed by non-compliance is CF, even in play.

Appendix 2: Coding Rules, Cont.

Notes on coding individual behaviours within CF

Commands

These are declarative statements and questions which directly require the other to alter their behaviour, or which demand a change in the speaker's state, such as demands and requests for permission. It is much more forceful than a suggestion in that it does not offer an option

- e.g., - "Please will you put your shoes on now"
- "I want a drink"

May take the form of a polite request, e.g., beginning with "Could you...?" or "Would you...?"

- e.g., - "Would you put the TV on?"
- "Could you get it all up please?"
- "Can I go upstairs?"

Prohibitions count as commands:

- e.g., - "Would you get off there"
- "Don't touch the biscuit"

Requests for permission count as commands. This is because the other person can comply or not to them, hence they often lead to CF.

- e.g., - "Can I go in the garden?"
- "Can I have a sweet?"

Thus "Can I have a sandwich?" plus stated refusal is CF.

Child Annoy

Repeat of previous (within visit) clearly bad or prohibited behaviour, e.g., "do not hit your sister," followed by sister being hit, clear destruction, clearly rude insult, threat, or shouting at someone. If not a previously prohibited act then it has to be more than just appear provocative. To be destructive, the act must cause more trouble than just need tidying up, i.e., it must ruin, waste, or potentially damage something:

- e.g., - deliberately smash plate
- deliberately scribble on walls
- deliberately break toy

Compliance

Compliance is coded for the sole purpose of defining conflict (see above).

Appendix 2: Coding Rules, Cont.

Compliance is coded when someone agrees to carry out what is requested of them in a command or suggestion, or when they do carry it out. In either case (the assent and the actual act of compliance), compliance is coded with no time limit. However, if the person agrees to do what is asked of them, but then at any time during that observation session clearly fails to carry it out, then non-comply is coded.

Deferred compliance is not non-compliance, e.g., “later,” “hang on,” “in a minute,” because the person is apparently saying “yes.” However, “not now” is non-compliance.

Non-compliance

Non-compliance is coded for the sole purpose of defining conflict (see above).

Coded when a person does not comply with commands or disagrees with it verbally. This includes not replying verbally or physically.

All refusals are non-compliance, even if command is unreasonable.

Outcome of conflict

During segments of CF make note of the outcome of conflict (i.e., who “won”), using F/M, C, or difficult to judge (N). Where CF evolves into more than one issue, the outcome of the issue that started the conflict should be noted and a second Outcome coding should be made at the end of the episode of CF. Where there is a “CF sandwich,” i.e., the same issue starts and finishes the episode, but there have been one or more other CF issues within the episode, the outcome of the penultimate issue should also be coded.

INDIVIDUAL CODING CATEGORIES

Commands / compliance

In general, the count of commands should tally with the count of child comply / non-comply. In cases where it is unclear whether commands by parent have been complied with or not (e.g., parent and child are out of view), commands should be counted but compliance and non-compliance should not be rated.

If a child changes behaviour as the result of a command and the rater is unsure if the child is complying or not, compliance should be rated in the absence of evidence to the contrary.

NB: When a command is in one time segment and the response is in the next, these should be counted as occurring within the same time segment, although if the response to a command in the first segment is a non-comply, this cannot be coded as conflict.

Appendix 2: Coding Rules, Cont.

Commands

Where a name is uttered in a commanding style immediately preceding any command, e.g., "George! Put your toys away!" code as only one command. Subsequent uses of child's name as a command within a time segment should be coded as vague commands.

Note both vague and non-vague (clear, specific, to the point, uncritical, age appropriate, *do*) commands.

Vague commands are:

- Commands phrased as a question
- Vague or implied, e.g., "I don't like you doing that," "Let's clear up"
- Use of child's name as a command
- Indirect, e.g., "Don't you think you ought to tidy up the toys now?" "Do you want to put the toys away now?"
- Descriptive, e.g., "Look, you're spilling your milk"
- Commands that are over-long (strings of more than two non-vague commands)
- Non-vague commands that are repeated more than two times in succession or repeated within time segment
- Commands in a string that begins with a vague command
- Commands made with explanatory rationale where there is no pause between command and rationale
- Commands disguised as a game
- Distractions / encouragements, e.g., "Come on," "look"
- Counting, e.g., "One, two, three" when used as a threatening command
- Commands starting with "Let's do..." (e.g., "this clearing up together")
- Obvious hint not expressed at index child
- Non-vague commands that are not age appropriate
- Non-vague commands issued in a clearly negative tone

verbosity

Code any utterance from parent between issuing a command and the child's response (whether complying or not) as verbosity within time segment, whether or not this command has led to CF, and regardless of the number of commands issued by the parent. If CF is initiated by child's demand or Annoy behaviour, code every parent utterance as verbosity after their initial response to the child. Stop coding verbosity when the conflict has been resolved.

Where parent command leads to compliance, but also leads to discussion or argument, e.g., "Why should I?" "Because I asked you to," code each parent utterance as verbosity until the discussion or argument relating to the issue in question ceases.

Appendix 2: Coding Rules, Cont.

Negative Behaviour

Code presence or absence of the following negative behaviours per segment:

Yell (Y) = 0 - 2 separate definite yells or angry tone (or combination) per segment

Yell+ (Y+) = 3 or more separate instances of, or continuous, definite yelling or angry tone (or combination) per segment

Rude (R) = swearing and/or rudeness and/or heavy sarcasm during segment

Rude+ (R+) = 3 or more separate instances of 'rudeness' as defined above

Derogatory remarks (R)

Aggression

Code presence or absence of aggression per segment, i.e., hit, shove, shake, actively reject, rough physical manipulation.

Aggression (AG) = 0-2 separate hits

Aggression+ (AG+) = 3 – 10 separate hits

Aggression++ (AG++) = more than 10 hits or continuous hitting

Threat

Count number of threats per segment. Threats are defined as aversive consequences for non-compliance to a command or prohibition. They must be consequences imposed by the speaker rather than natural consequences of the behaviour itself, e.g., "Stop it or you'll get a smack," "Stop it or else" (implied consequence). Where the threat is also a command, double rate under Command and Threat.

Follow Through

If the parent issues warnings, reminders, bargains or threats, and subsequently "follows through" on these strategies, the original code should be crossed through to indicate this. In general, "follow through" can be coded provided this has occurred before the end of the visit.

Beg / Coax

Make note of use of begging or coaxing. This can take the form of a command or praise, e.g., "Come on!" "Please!" "That's it," "There's a good boy."

Coding Grid

Sheet Number.....

[illegible]

APPENDIX 4

Consent Form

University of Oxford

Department of Applied Social Studies and Social Research



BARNETT HOUSE
WELLINGTON SQUARE
OXFORD OX1 2ER
TEL 01865 270325
FAX 01865 270324
<http://barnett.apsoc.ox.ac.uk>

AQREC Study No: A99.044

Research Group Telephone No:
01865 270320

CONSENT FORM

RESEARCH STUDY COMPARING TWO COMMUNITY PARENTING PROGRAMMES

Researchers:

Frances Gardner
Kathy Sylva
Jenny Burton
Sue Kirkpatrick

Please initial box

1. I confirm that I have read and understood the Information Sheet dated for the above study and have had the opportunity to ask questions. ()
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without affecting the help I receive or my legal rights being affected. ()
3. I agree to take part in the above study and understand that this will involve video-recording during part of the visit. ()

Name of participant

Date

Signature

Name of person taking consent
(if different from researcher)

Date

Signature

Researcher

Date

Signature

1 copy for participant; 1 copy for researcher

APPENDIX 5
Letter Indicating Ethical Approval



APPLIED & QUALITATIVE
(Formerly NAPREC)
RESEARCH ETHICS COMMITTEE

Manor House
Headley Way, Headington
Oxford OX3 9DZ

Tel: 01865 222758
Fax: 01865 222699

Our Ref. JB/LAB/A99.044

2nd September 1999

Dr Frances Gardner
Dept of Applied Social Studies & Social Research
Barnett House
Wellington Square
Oxford OX1 2ER

Dear Dr Gardner

RE: A99.044 - Educating parents of hard-to-manage children: an evaluation of community parenting programmes

Thank you for your letter dated 25th August 1999, letting me have the further details on this project. I am happy to confirm ethical approval, and wish you every success with the study. I would be very grateful if you could send me a copy of any publication that may arise from this study.

May I remind you that if the investigators do not follow the protocol, or make changes to the protocol, without informing AQREC then Ethics Committee approval will be withdrawn. In addition AQREC should be made aware of any adverse events.

AQREC final approval is contingent on the appropriate indemnity.

Yours sincerely

Corrain Bushy.

Dr Jenny Butler
Chair
Applied and Qualitative Research Ethics committee

continued

Chair: Dr Jenny Butler

The Oxford Radcliffe Hospital
A National Health Service Trust

APPENDIX 6

Key Characteristics of Variables for Reliability Analysis (N=15)

Variable	Mean	Minimum	Maximum	Standard Deviation
Researcher 1				
Laxness*	-17.13	-0.61	0	19.31
Over-reactivity**	1.73	0	9.00	2.49
Verbosity***	10.53	0	34.00	11.17
Compliance	6.87	0	23.00	7.62
Non-compliance	9.33	0	34.00	11.33
Total Number of Commands	16.53	0	57.00	18.80
Non-vague Commands	6.93	0	28.00	9.07
Vague Commands	9.60	0	31.00	10.39
verbosity	1.07	0	3.00	1.16
Yelling (Yell + Yell+)	0.47	0	5.00	1.36
Rudeness + Aggression	0.20	0	2.00	0.56
Begging/Coaxing	0.33	0	4.00	1.05
Conflict	3.20	0	10.00	3.28
Conflict Outcome ('parent wins conflict' minus 'child wins conflict')	0.93	0	4.00	1.28
Researcher 2				
Laxness*	-16.20	-47.00	0	16.20
Over-reactivity**	2.93	0	11.00	3.73
Verbosity***	10.93	0	43.00	12.27
Compliance	7.20	0	23.00	6.92
Non-compliance	7.87	0	31.00	9.90
Total Number of Commands	15.53	0	47.00	15.59
Non-vague Commands	6.73	0	31.00	8.14
Vague Commands	8.80	0	32.00	9.49
verbosity	2.13	0	11.00	3.04
Yelling (Yell + Yell+)	0.67	0	5.00	1.45
Rudeness + Aggression	0.13	0	2.00	0.52
Begging/Coaxing	0.20	0	3.00	0.77
Conflict	3.00	0	10.00	3.40
Conflict Outcome ('parent wins conflict' minus 'child wins conflict')	0.87	0	4.00	1.19

*Laxness = Begging/Coaxing – Total Number of Commands – Outcome

**Over-reactivity = verbosity + Yelling + Rudeness + Aggression

***Verbosity = Vague Commands + verbosity

APPENDIX 7

Key Characteristics of Variables for Correlational Analysis Involving Observed and Parent-Reported Parenting Variables

Observed Variables

Variable	Mean	Minimum	Maximum	Standard Deviation	Number of Parents for whom Behaviour Observed
Verbosity (vague commands + <u>v</u> erbosity)	44.33	1.00	170.00	37.53	40
Total Number of Commands	74.10	5.00	262.00	56.59	40
<u>v</u> erbosity	5.55	0	16.00	4.30	37
Yelling (Yell + Yell+)	3.85	0	17.00	4.53	25
Rudeness	0.90	0	4.00	1.32	16
Begging/Coaxing	0.88	0	5.00	1.34	17
Conflict	16.33	1	45.00	10.87	40
Conflict Outcome (‘parent wins conflict’ minus ‘child wins conflict’)	3.70	-2.00	11.00	4.03	40

(Continued)

APPENDIX 7, Cont.: Key Characteristics of Variables for Correlational Analysis
Involving Observed and Parent-Reported Parenting Variables

Item	Content of Item	Mean	Minimum	Maximum	Standard Deviation
4	Amount said when telling child not to do something	4.63	1.00	7.00	2.23
6	Arguing with the child	4.13	1.00	7.00	2.05
8	Limit setting	1.97	1.00	7.00	1.30
9	Amount said when child misbehaves	3.05	1.00	7.00	1.97
10	Raised voice	5.45	1.00	7.00	1.74
12	Firmness of command	2.84	1.00	7.00	1.81
19	Keeping to limits set	3.98	1.00	7.00	1.99
25	Use of 'bad language'	2.59	1.00	7.00	1.80
26	Keeping to limits set	2.68	1.00	7.00	1.65
28	Use of insulting terms aimed at child	2.25	1.00	6.00	1.56
29	Verbal response to complaints from child	2.90	1.00	7.00	1.72
30	Keeping to limits set	2.58	1.00	7.00	1.93